

We thank the reviewers for their very detailed comments, which prompted us to reflect deeply on several of the points raised. We hope that our responses are sufficiently precise and well-supported. The original comments are in bold font, and the responses in normal font.

### **General Comments:**

**The manuscript presents a novel approach to obtaining U-Pb ages of carbonates using isotopic maps and guided by statistical approaches to obtain the best age in what can be considered virtual spots. This is an interesting approach, and we can see the benefits the approach can have. The authors present a large amount of data from a number of samples (some of which have been previously dated for comparison).**

### **Specific Comments:**

- In the introduction, the manuscript heavily relies on references and comparison with other studies, which when one isn't familiar with all of them makes reading and following the arguments a bit difficult. Particularly as the reader doesn't know the details of the approach this paper takes at that point in the text. Therefore, we wonder if some of this might be better suited for the discussion instead of the introduction. We would welcome it if the abstract and the introduction would focus a bit more on the rationale for using this approach. Why is this new method needed, what is the overall problem, that justifies using the approach the authors present? The text says that the ages are similar to the traditional approach but that the uncertainties can be worse. So, the authors should be clearer what the advantage of this method over the other methods is.**

In the introduction, we will aim to clarify the relevance of using isotopic imaging for geochronology, as well as the need to develop approaches that yield the most reliable ages possible. However, we do not intend to engage in a comparison of 'which data processing model is the best.' In this contribution, we present an alternative approach to those already published. We do not wish to portray our approach as a breakthrough that renders previous work obsolete. On the contrary, our method was developed through reflections inspired by existing approaches. As stated in the introduction and the discussion, each method has its own strengths and weaknesses, and ours is no exception. Through this contribution, above all we aim to share with the community the progress of our ongoing work on LA-ICP-MS map processing for U-Pb geochronology.

**After stating here in the methods that some samples have been treated differently none of this is mentioned afterwards in the results of discussion. Even if it doesn't make a difference. It would be important to mention that explicitly. If it does make a difference could some of the results be influenced by this and if yes how?**

We suppose you give details of the different treatments in the following lines.

- Line 128: Why was the laser frequency changed from 100 to 500 Hz, please explain this change and the advantage of using one over the other. Additionally, can you detail if and what change this has on the results?**

In this contribution, we draw on analytical results obtained over the course of several years and across numerous research projects. By default, we work at 500 Hz on thick samples (mounts) to maximize signal intensity. With this setup, we typically obtain 100k–200k cps for  $^{238}\text{U}$  on the NIST SRM 614. The work performed at 100 Hz (ARB20-2D) was conducted on a thin section. The only reason for this was the shallower ablation depth, which prevents drilling through the section. It is of course associated with lower number of counts (70–100k cps). This change in repetition rate has no impact on how the data are processed using our method.

- Line 137: For sample BH14 only Ar was used as a carrier gas. It has been shown by Eggins et al., 1998 that Ar and He have quite different transport qualities. How did the use of only Ar as a carrier gas influence the results?**

We cannot answer. BH14 was one of the very first samples analysed using imaging in our laboratory. We initially tested Ar to assess the feasibility of carbonate imaging, knowing that the washout time would be longer. Later, switching to He was motivated by its faster washout time, which provided greater responsiveness, although it also led to slightly more signal variability (based on standards WC1 and NIST), without any significant impact on the quality of the results in terms of precision and accuracy. We did not repeat tests using Ar after the analyses of BH14.

- Line 149: Why was NIST SRM 614 used after 2020 and 612 before 2020? Are there any differences in the results or uncertainties?**

Initially, we worked with NIST SRM 612 for two main reasons: (i) it was already used to tune the spectrometer at the beginning of each session, making it a practical choice. Our custom-built ablation cell does not allow for multiple standards alongside the unknown sample, so keeping NIST SRM 612 saved time; and (ii) tests using NIST SRM 614 revealed slight inhomogeneities in  $^{206}\text{Pb}/^{238}\text{U}$  ratios (as has been reported elsewhere). However, since the concentrations in NIST SRM 614 are closer to those found in natural calcite, its use for calibration appears more appropriate

and justifies the switch. Objectively, we did not observe any significant differences depending on the standard used. It is also worth noting that some research groups use NIST SRM 612 for Pb/Pb ratio correction (e.g., Parrish et al., 2018).

- **Use of rainbow colour maps. Jet or rainbow colour maps have been shown to be not a good choice. Both in terms of accessibility (colour blind and other sight impairments) the rainbow scheme is also misleading normal normal-sighted people due to a higher sensitivity of the human eye for certain colours which leads to a visual bias. Therefore, it should not be used:**

<https://blogs.egu.eu/divisions/gd/2017/08/23/the-rainbow-colour-map/>

**Have a look here for alternative suggestions**  
<https://www.fabiocrameri.ch/colourmaps/>

Thank you for your suggestion. We will revise the color schemes in the figures to ensure they are accessible to visually impaired readers.

- **The comparison of the ages already determined by other studies (AUG-B6, BH14, DBT, Senz7) and this manuscript isn't very clear. In section 5.1.1. it is mentioned, but in none of the tables or figures (despite a reference to it) is it shown clearly.**

The reference ages are provided in Table 2 along with the appropriate citations, as well as in Section 2. In Section 5.1.1, Table 2 is indeed not referenced, which we will correct.

- **A lot of the ages are quoted without external error propagation. How much would external uncertainty add? I assume a good amount of external uncertainty comes from the standards? What other sources are there and what is the justification to ignore them? When comparing ages from previous studies with this one external error propagation would be needed to understand the full extent of how they compare to each other.**

We suppose that what you call 'external error' is that coming from the standards (referred to as "External  $2\sigma$  err req'd (each pt)" in Isoplot). In our current procedure this additional uncertainty is directly calculated by Lolite4 from NIST SRM 614, based on the paper of Paton et al 2010 (reference will be added to the Methods section for clarity). Note that since we follow closely the recommendations of Horstwood et al. (2016), such excess variance is added directly onto the data points (ellipses) and not onto the final age. Additional systematic uncertainty includes a long-term variance (2%) that should allow confident comparison with ages obtained by other studies.

- **In section 5.2 the choice of data processing approach shows large (up to 50 Ma) variations in age (shown in figure 6 as well). How should a user choose which one to use? In the text, the authors mention that they chose the one closer to the expected age. But what if one doesn't know the expected age? Wouldn't this also mean that people might reject the 'true' age because they didn't expect it?**

Thank you for your pertinent observation. We fully agree that obtaining reliable ages from low-concentration samples—such as C6-265-D5—remains one of the main limitations of the approach presented in this study. It is indeed essential to verify whether a given sample is suitable for accurate U-Pb dating using this method. We address this issue in detail in our response to Don Davis, but we will also clarify it further in the revised manuscript. Specifically, we propose including a more explicit discussion outlining the importance of systematically comparing ages obtained from both Tera-Wasserburg and Wetherill regressions, in addition to evaluating conventional statistical parameters such as the MSWD. If the TW and Wetherill concordia ages do not overlap within their respective uncertainties, this should be taken as a strong indicator of potential bias and sample unreliability. Our tests suggest that, in such cases, averaging (smoothing) the mass scans can help reduce the impact of erratic isotope ratios, leading to better agreement between TW and Wetherill ages. However, this comes at the cost of reduced precision due to the lower number of pixels per virtual spot. Whether this trade-off improves accuracy remains an open question. The example of C6-265-D5 is promising, as the age obtained after smoothing aligns well with independent geological constraints. That said, we acknowledge that further work is needed to assess the robustness of this approach across a broader range of low-concentration samples. We will add this discussion to the revised manuscript, both to clarify the limitations and to outline potential future improvements.

- **In section 5.3.1 line 348 the authors refer to a map of sample Cot02a that has more precise ages, but then say it isn't presented here? Why is that? Why not at least provide this in the supplement?**

We did not include the second map (produced during the same analytical session) in the main manuscript because our objective here was to focus on the potential of isotopic mapping to distinguish and date multiple generations of calcite—specifically, matrix versus fracture domains. The second map simply does not contain any fracturing and therefore does not contribute directly to this particular discussion. However, we agree that providing it may offer useful context, and we will include this additional map in the supplement.

- **In lines 427-428: The sentence 'In the case of reliable samples, it is expected that the position of the spots does not influence the calculated ages.' What would you consider a reliable sample?**

Thank you for pointing this out. We agree that the definition of a “reliable sample” should be clarified. In the revised manuscript, we will specify that by “reliable sample,” we refer to a sample with a homogeneous age and common Pb composition, sufficiently high U and Pb concentrations, and a wide enough range in isotopic ratios to allow for both accurate and precise U-Pb dating. We will revise the phrasing accordingly in the relevant section.

**•What if your sample isn’t reliable, can you still use the approach?**

In our view, the proposed approach is particularly useful for assessing the reliability of a sample for geochronological purposes. As a first step, it is important to verify the consistency of the weighted mean ages obtained across the dataset. To support this, we propose to include a supplementary figure showing the weighted mean ages calculated for the different samples, based on 100  $\mu\text{m}$  x 100  $\mu\text{m}$  virtual spots in both the Tera-Wasserburg and Wetherill spaces. However, consistency alone is not sufficient to guarantee accuracy. As demonstrated by the C6-265-D5 example, coherent age estimates can still be biased. Therefore, as a second step, we believe it is essential to compare TW and Wetherill ages, which should agree within uncertainty. As discussed above, a discrepancy between these two regressions indicates potential bias and questions the reliability of the sample. We will expand the discussion on these points in the section addressing the limitations of our approach. Please also see below for additional details.

**•We would appreciate it if the authors could provide a bit more information on the limitations of this method. Particularly as carbonates can be complex in their formation and thereby age pattern. We would also appreciate it if the authors would touch on potential user biases that could be introduced when choosing one approach over another. It has been shown that data reduction and the choice of approach can make a difference in the final result. Particularly here with the possibility of choosing many virtual spots and statistical approaches, some clear guidelines for the reader would be helpful.**

In our virtual spot approach, we propose several strategies to obtain ages: the moving grid, sub-image, and Rectis methods. Among these, the moving grid is the primary and recommended approach and should always be used as the first step. We will clarify this point in the discussion. The sub-image and Rectis methods are complementary but optional. Specifically, the sub-image method presented here is more of a proof of concept, demonstrating the potential to date very small areas using our approach. Nonetheless, it can also contribute additional age constraints for larger samples through a weighted mean of sub-image results.

To help guide users through the most appropriate use of our method and to mitigate its main limitation—namely, the risk of biased ages for challenging

samples—we will add a workflow to the discussion. This will be presented both in text and as a graphic and will outline a logical sequence of steps:

1. Initial age calculation: Calculate a series of ages using 100  $\mu\text{m}$  x 100  $\mu\text{m}$  virtual spots in both TW and Wetherill spaces.
2. Check consistency: Assess (i) internal consistency of the ages in each space, and (ii) agreement between TW and Wetherill ages. Both conditions must be satisfied. If not, investigate the potential causes (e.g., low U/Pb concentrations, insufficient isotopic spread, multiple age populations). The datability of the sample should be questioned.
3. Age selection:
  - (a) If the ages are consistent, select the most statistically robust result (e.g., lowest MSWD, highest precision), potentially adjusting the virtual spot size to improve the result.
  - (b) If the TW and Wetherill ages are inconsistent, recalculate the ages after averaging the mass scans. If the TW and Wetherill results converge, return to step 3a. If not, the sample is likely undatable.
4. Optional steps:
  - Apply the Rectis method to obtain a potentially more precise age or one with better statistical parameters.
  - Use the sub-image method as an additional test. The weighted mean of the sub-image ages should ideally match the result obtained using the moving grid.

This workflow will be added to the revised manuscript to help users apply our method more effectively and transparently.

**•A number of studies (e.g., van Elteren et al., 2018, Norris et al., 2021) have shown that aliasing and other artefacts can be created by LA-ICP-MS mapping. Have you observed any such effects, how are you mitigating such effects and how would this influence your ages?**

We have not encountered such issues, except for a slight smearing effect related to the washout time, which should have no significant impact on the accuracy of the obtained ages (see also our response to Reviewer 1). However, it is plausible that effects such as aliasing could have a more substantial influence on the calculated isotopic ratios and, consequently, on the ages derived using our approach. This remains to be tested and represents an avenue for future investigation.

**Technical Comments:**

- **Use of the word 'image' when referring to the laser map, in many cases even using 'image map'. Images imply that pixels are acquired simultaneously, which is not the case for LA-ICP-MS. Therefore, the word map should be used in this case. Please change this throughout the manuscript.**

OK. We will follow these recommendations.

**Line 55 & 62: use  $^{238}\text{U}/^{208}\text{Pb}$ , we were wondering if this is a typo and should say  $^{238}\text{U}/^{206}\text{Pb}$ .**

This is not a typo. The  $^{238}\text{U}/^{208}\text{Pb}$  ratio is used in the approach of Drost et al. (2018). See their paper for additional detail.

**Line 76/77: 'Both allowed to obtain highly spatially resolved image maps (25  $\mu\text{m}$  rasters) and with a good analytical sensitivity.' What is good analytical sensitivity in this case, can you please quantify this?**

Taking NIST SRM 614 as an example, at 500 Hz, analytical conditions used give about 100 kcps – 200 ckps / ppm  $^{238}\text{U}$  depending on the session. These values will be added to the text. Limits of detection and quantification are provided I. 147-148.

- **Line 169: 'This script as well as the ones described below are publicly available (Hoareau et al., 2024).' Instead of saying it is publicly available and referring to another paper it should say where the code is available. Good practice is to publish the code on GitHub and then provide the link here. Make it easy for people to find and use!**

We will ensure that the code made available on Zenodo is cited appropriately, as indicated in the Data Availability section. At this stage, we do not plan to publish the code on GitHub.

- **Line 172 should say virtual spot**

Yes.

- **Inconsistent use of NIST glass names, sometimes NIST SRM 612/614 other times NIST 612/614. We suggest to always use the same labelling.**

OK. We will use NIST SRM 612 and NIST SRM 614.

- **The use of the word spot height is a bit confusing, we suggest using width instead when talking about the vertical extent and then**

**speaking about line length when speaking about the horizontal extent.**

We had the same remark from first reviewer. It will be changed.

- **Line 183-185: “To achieve this, it may be necessary to adjust the size of the virtual spots very slightly (e.g., 51 µm x 50 µm instead of 50 µm x 50 µm) to ensure an integer number of pixels per spot.”**
- **Why is it, is it because of the Python language/computing characteristics?**

The user interface allows the user to specify virtual spot sizes in microns. The calculation then divides this length by the chosen mass scan duration to determine the number of pixels, which in turn defines the tolerated offsets between successive moving grids. Slightly adjusting the virtual spot size can result in a greater number of possible grid positions, as it leads to more integer values in the calculations. This behaviour is independent of the programming language used.

- **Fig 1. The use of the arrows from A to B&C is confusing as they are not derived from A. They simply show a different configuration or the grid. Therefore, we suggest deleting the arrows. Furthermore, it would be good to provide the information on what map is shown. Is that a raw map? What element/ratio does it show? What sample is it?**

We agree. We will remove the arrows and add the ratio displayed ( $^{238}\text{U}/^{206}\text{Pb}$ ). Sample is BH14 as explained in the caption.

- **Fig 2. This figure is quite busy, are all images needed? Could some of them go into the supplement?**

Since the Figure presents the maps used and discussed in the study, we feel that keeping them in the main text is a minimum.

- **In the caption of Fig 2, it says that concentrations were estimated from NIST SRM. First of all, which NIST was used for which sample? And secondly, what do you mean by estimates? Why not do a proper calibration of them?**

Concentrations are semi-quantitative estimates based on analysis of NIST SRM 612 (before 2020) and NIST SRM 614 (after 2020), as calculated from Iolite4. Purely quantitative concentrations would require the analysis either of  $^{42}\text{Ca}$  mass or that of a carbonate standard of homogenous U, Th and/or Pb concentration. The term ‘estimates’ can be replaced by ‘semi-quantitative’.



- **Fig. 7: The maps at the top have two areas highlighted that represent two calcite generations. What is the rest of the map? Nowhere does it say what it is and based on what the areas highlighted have been chosen.**

The rest of the map is made of quartz (in blue) and calcite (in red). The highlighted areas have been chosen based in a sufficiently high  $^{238}\text{U}/^{206}\text{Pb}$  for calcite rhombs (area 1) and the presence of a calcite-filled fracture (area 2). We will label the quartz grains and non-analysed calcite, as well as provide more detail in the text.

- **All isotope maps would benefit from a scale bar**

Right. We need to add them.

- **The labelling of the elements and isotope ratios next to the maps in the figures is very small and not well-readable. Suggest increasing the font size.**

OK.

**This is a co-review from Barbara Kunz & Igor Figueiredo**