

Reviewer 1

Many thanks to Referee #1 for their insightful comments and helpful feedback. Our responses to their queries and revisions (bold) are listed below. Line numbers reference the original submission.

Specific comments

- 1. Lines: 213-214: Could there not be an issue with heterogeneous beta dosing when using this approach due to attenuation? So that 0-1 mm grains received more radiation compared to the 1-2 mm grains?**

That is absolutely a possibility and likely that the grains in the outer 0-1 mm received a higher dose rate than those at the 1-2 mm depth grains given that the surrounding sediment matrix has a dose rate of 1-2 Gy higher than the limestone clasts themselves. Our method of sediment extraction from the limestones could not be achieved at a higher precision than +/-1 mm as estimated with callipers, so we have to accept this additional source of uncertainty. This is why our calculated dose rate for these grains was taken to be an average of the dose rate modelled for the outer 2 mm of the limestone rocks (Section 4.3.4, Fig. 6).

- 2. 225-226: How sure are you that this assumption is correct? Is it not quite possible that the different surfaces have had different exposure/burial histories? If you didn't measure any luminescence-depth profiles, how do you know that the signal was bleached? Would it not have made sense to try to measure signal-depth profiles?**

The special approach we took to process the limestone clasts certainly have limitations, which is why we felt it necessary to be transparent about the assumptions they require. We agree that the assumptions that our methods require may not be correct, and this is further discussed in the results and discussion of the study (Sections 5.1.1-5.1.3). We also would expect that the rock surfaces would have had varying exposure histories as clearly shown by the volcanic clasts from Cave Valley. Unfortunately, the limited material within the limestone clasts made it unrealistic to extract detrital sediments from core slices of limited diameter (as one would do to obtain luminescence-depth profiles), so measurements were made from the near-surface detrital grains from the entire rock surface. Single-grain measurements (as opposed to measurements from rock slices) were made to help identify grains that were most recently bleached (using MDM), and therefore likely more representative of the final burial age of the rocks, even if they had uneven light exposure prior to burial (Section 5.1.3).

3. 244-245 What was your rationale for crushing the slices instead of measuring them whole on the carousel? Would you not have avoided the issue of weak signals if you had just measured the slices directly?

We chose to crush the rock slices for two reasons: i) our beta radiation sources of our readers were not calibrated for slices, but rather only sediments, and ii) the volcanic rocks that we measured crumbled easily during slicing; most could not be measured as intact slices that would sit in a slot on the reader carousel.

4. 253-254: My comment here is similar to that for the limestone: what if the surfaces have different histories? Based on Figs 12 and 13, it seems that for some rocks, these De values seem comparable, but in others, not at all. Would it not have made more sense to measure the surfaces separately?

We cored through the entire thickness of the volcanic rocks (data shown in Figs. 12-14), so, for the rocks where we could record and label the top side and the bottom side (Rocks 7, 11, 13, and 18), the data shows the age-depth profile for both sides. These data show that most rock surfaces were not evenly bleached on all sides (Section 4.4.8).

5. 782-783: How did you calculate this bleaching rate (>10 mm in less than 200 years)? How would this translate to the much slower bleaching pIR290, since this seems to be the signal that provides ages from the centre of the clasts that agree with age control?

This was an estimate inferred from bleaching rates calculated by past researchers (Ou et al., 2018, Lehmann et al., 2018, etc. mentioned in the same paragraph), however bleaching tests were not performed on our samples and precise bleaching rates for our samples were not obtained. To avoid speculation and confusion, we have removed that statement. We agree that the bleaching rate of the PIRIR290 signal would be much slower than that of the IR50 signal, but are open to the idea that in gravel sized rocks that are small enough, the PIRIR290 signal may have been depleted throughout the entire rock thickness in the lakeshore environment prior to final burial (lines 783-785).

6. 784-785: Intriguingly, you seem to get the best ages from the pIR290 signal from the centre of your clasts, since it would mean that you have had an incredibly rapid bleaching of the pIR290 signal. Have you done any bleaching tests to assess how quickly the pIR290 might bleach in your samples?

No, unfortunately we were unable to perform bleaching tests, but agree that bleaching processes for gravel sized rocks in pluvial lake nearshore environments should be investigated (see lines 831-833). The question is, despite the relatively slow bleaching rates of the PIRIR290 signal (compared to the IR50 signal), could the

frequency and duration of bleaching events deplete this signal throughout the entire rock thickness prior to final burial.

Technical corrections

7. Table 4: I'd prefer if the DRT ratios were included here instead of Passed/not passed

Table 4 DRT ratios will be updated as follows:

Rock 1 -> SG DRT ratio = 0.97 ± 0.07 n=10

Rock 2 -> MG DRT ratio = 0.98 ± 0.02 , n=4; SG DRT ratio = 0.99 ± 0.02 , n=285

Rock 5 -> MG DRT ratio = 0.92 ± 0.04 , n=3

Rock 9 -> MG DRT ratio = 1.00 ± 0.04 , n=6; SG DRT ratio = 0.97 ± 0.04 , n=7

Rock 10 -> SG DRT ratio = 0.91 ± 0.03 , n=20

Rock 11 -> MG DRT ratio = 1.07 ± 0.04 , n=4; SG DRT ratio = 0.99 ± 0.04 , n=9

Rock 18 -> SG DRT ratio = 0.96 ± 0.03 , n=53

8. Section 4.3.6 and 4.4.7: Please report the sigma_b values used for MAM.

Sigma b values used for MDM will be included in Table 6 as follows:

Rock 2 (90-125 μm) -> 0.20

Rock 2 (63-90 μm) -> 0.20

Rock 10 -> 0.20

Rock 18 -> 0.17

9. Lines 545-546: Are you referring to the feldspar grain K content in such rocks or the total rock K content here? I think this sentence can be improved for clarification.

We can clarify the statement to say that our total dose rates assume an internal K content of $10 \pm 2\%$ for our measured grains following Smedley et al. (2012). This is to acknowledge that the "grains" we've measured from our volcanic rocks may actually be clumps of many finer mineral grains that are a mixture of minerals, as discussed in the following paragraph. It is likely that our PIRIR signals emanate from a range of feldspar types.

Lines 544-546 have been revised to:

"Total dose rates include an internal beta dose rate component that assumes an internal K content of $10 \pm 2\%$ for our measured grains following Smedley et al.

(2012), but we acknowledge that this value may be high as K contents of more intermediate or andesitic rocks, dominated by plagioclase, are typically lower than 10% (Taylor, 1968) and there is uncertainty in the size of the measured feldspar grains in our samples (see below).”

10. 556-562: Is this a test performed on one rock to check the impact of internal K-content? Or did you use this approach on all rocks? Please clarify.

As stated in line 557, we performed the calculation from rock slices from Rock 4 only. We’ve revised line 556 to read:

“Given the uncertainties regarding feldspar internal K content and the grain size of volcanic feldspar grains in our study, a test was conducted to determine the sensitivity of our age calculations to variations in these parameters. Dose rates and fading-corrected ages were re-calculated for all rock slices from Rock 4 assuming:...”

11. 612-614: I can’t fully understand this sentence. With heterogeneous light penetration, do you mean that one surface was exposed longer?

We agree this statement should be clarified. The statement:

“Most luminescence De profiles show a decline of one or both signal De values toward at least one surface of the rock indicating limited, and heterogeneous light penetration into the rock surfaces. Where De values rise with depth, pIRIR290 De values increase at a more rapid rate than the IR50 De values, and to a higher apparent saturated level, leading to vertically enhanced versions of the IR50 De profiles; this is typical of combined IR50 and pIRIR luminescence-depth profiles reported elsewhere (e.g., Sohbaty et al., 2015; Freiesleben et al., 2015; Jenkins et al., 2018).”

will be revised as follows:

“Several rock PIRIR De profiles (Rock 4, Rock 7 core 1, Rock 11, Rock 13, Rock 18 cores 1 & 2) rise with depth into the rock at either the top or bottom side. For example, the Rock 7, core 1 PIRIR De profile starts at its lowest point at the bottom surface of the rock (left), then rises with depth into the rock (Fig. S24B). The IR50 De profiles of the same rocks typically rise in a similar pattern, but at a much lower rate, yielding much more subdued (flatter) profiles. Similar patterns have been observed in IR50 and PIRIR luminescence-depth profiles reported elsewhere (e.g., Sohbaty et al., 2015; Freiesleben et al., 2015; Jenkins et al., 2018).”

Reviewer 2

We thank Referee #2 for their thoughtful comments and constructive feedback (in bold below). Our responses and revisions are outlined here.

- 1. The study rightly highlights the important issue of inhomogeneous bleaching history, which is a crucial aspect that could be further emphasized in the abstract. It is clear that one cannot simply assume that dating the entire rock surface yields an accurate age. The SG approach is a valuable and effective method for addressing this challenge.**

We have highlighted this aspect in the abstract, starting at Line 19.

“Directly dating when these gravel clasts were last exposed to sunlight via luminescence is ideal but their limestone and volcanic lithologies prove challenging. Initial measurements from these lithologies show that feldspar luminescence signals are suited to single-aliquot regenerative (SAR) dose measurement protocols and show evidence for heterogeneous bleaching of rock surfaces. Polymineral extracts from dissolved limestone clast surfaces from Coal Valley that contain sufficient detrital sediment exhibited infrared (IR) signals with low to moderate fading rates. Single-grain ages from this detrital sediment, calculated using the minimum dose model, straddle the C-14 age estimate of the Pluvial Lake Coal highstand with one age consistent with the C-14 at 1σ .”

- 2. It is a minor limitation that no measurements were taken deeper within the volcanic rocks or from similarly larger samples, which would have provided useful confirmation of whether the luminescence signal is truly saturated or fully bleached before last burial. Additionally, lab-to-field saturation ratios could also offer valuable insight into fading effects.**

Measurements were indeed taken at depth (throughout the entire thickness) of the volcanic rocks, as well as from their surfaces (see description of both approaches in Section 3.2.2 and results in Sections 4.4.5 and 4.4.6).

- 3. Furthermore, I would have welcomed measurements from both rock types of samples exposed to light at the time of collection, as these would offer further insightful information on the bleaching processes.**

We agree that more measurements from modern, or sun-exposed rocks of similar lithology would be informative. Attempts were made to measure limestone samples collected at the ground surface from Coal Valley (Rock 2M and Rock 3M, Table 4), however they did not yield material with a luminescence signal to measure. As we state in the Conclusions and Abstract of the study, future research should include investigation of gravel bleaching processes in pluvial lake environments.

- 4. L 203: “for only for the limestone”**

Corrected.

- 5. 206: Traditional? There are several "traditional" preparation methods, such as slicing and grinding, but the choice largely depends on the rock type. You**

might want to reconsider or clarify the citation here to ensure it accurately reflects the context or specific method being discussed.

We have clarified the statement to refer to the common approach involving extraction of drill cores that are sliced (intact or ground and sieved) following the approaches of Jenkins et al. (2018) and Gliganic et al. (2019).

6. 212: 1. The outer secondary carbonate coatings were filed away with a file or Stylo-style Dremel tool. Could you specify how much material was removed (in mm)?

The coatings were thin (<5 mm thickness) and could be visually differentiated from the parent rock. They did not cover the entire rock surface, however, often just one or both top and bottom sides, or their coverage was patchy. These were removed manually, but their thickness was too variable to be measured. We've included a description of the coatings at line 212.

7. L- 223: During beach ridge formation, light penetrated the outer 2 mm or more of the limestone surface to bleach the signals from detrital quartz and feldspar minerals.

The assumption here treats the RSLD sample as if it were sediment, relying solely on surface bleaching. However, since this is a rock surface, there is an opportunity to extract more information by analyzing signal variation with depth. Do you have inner material or a luminescence-depth profile that could support or challenge the assumption of surface bleaching? Otherwise, the unique potential of RSLD compared to sediment may not be fully utilized.

The special approach we took to process the limestone clasts certainly have limitations, which is why we felt it necessary to be transparent about the assumptions they require. Unfortunately, the limited siliceous material within the limestone clasts made it unrealistic to extract detrital sediments from ~1 mm thick core slices of ~8-10 mm diameter (as one would do to obtain luminescence-depth profiles), so measurements were made from the near-surface detrital grains from the entire rock surface. As the referee as noted, single-grain measurements (as opposed to measurements from rock slices) were made to help identify grains that were most recently bleached (using MDM), and therefore likely more representative of the final burial age of the rocks, even if they had uneven light exposure prior to burial (Section 5.1.3).

The potential benefit of dating the limestone clasts (rather than the beach ridge sand matrix) in this context is that the clasts will have been much less prone to remobilisation and translocation after beach ridge formation than sediments.

- 8. 244: The polymineral slices were subsequently crushed gently by hand using an agate mortar and pestle and sieved into distinct grain size fractions between 125 and 250 μm for measurement.**

While it makes sense to extract known grain size fractions, did you assess whether the mechanical crushing process alters the luminescence signal, for example through induced sensitivity changes or signal resetting?

We did not assess this; crushing was done gently by hand using a mortar and pestle following the approaches of Brill et al. (2021, Earth Surf. Dyn.) and Gliganic et al. (2024, Quat. Geochron.). We agree that such tests would be informative, but we note Meyer et al. (2018, Radiat. Meas.) found that variabilities in luminescence-depth profile shapes does not seem to be dependent on sample preparation technique (i.e., intact, crushed or crushed and density separated slices, their Figure 1).

- 9. 278: Aliquots/grains were rejected from further analysis if the recycling ratio was beyond 10% of unity and if recuperation was greater than 5% of the sensitivity-corrected natural signal.**

Have you assessed whether the rejection of these data points introduces any bias or significantly alters the results?

We did not explicitly test the impacts of individual aliquot/grain rejection in our original study. A cursory look at the impact of recuperation and recycling ratio rejection criteria on our single-grain and multi-grain aliquot data are shown here in Figure A.

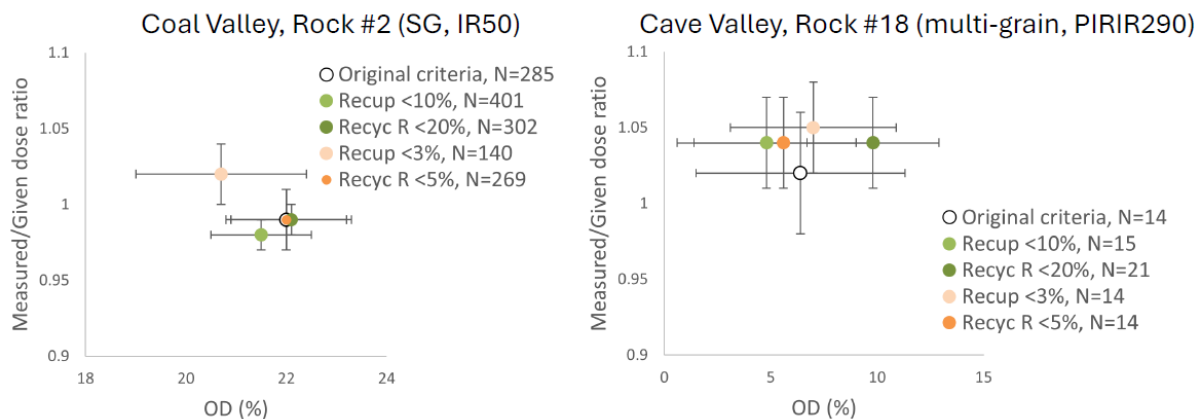


Figure A. Dose recovery test results (measured/given dose versus overdispersion) for limestone Rock #2 from Coal Valley (SG data measured with the IR50 signal), and volcanic Rock #18 from Cave Valley (multi-grain aliquot data measured with the PIRIR290 signal). Dose recovery test results have been re-analysed after changing either the recuperation rejection criteria or the recycling ratio rejection criteria. More lenient values are shown in the green, and more conservative or restrictive values are in orange. ‘N’ refers to the number of single grains/aliquots that were accepted after analysis. The

original study (hollow symbol) accepted aliquots or grains with recuperation <5% of the natural signal and recycling ratios <10% of unity.

Here, we re-analyse dose recovery test results using maximum allowable recuperation values of 3% and 10% while keeping the maximum allowable recycling ratio at 10%. Then we re-analysed the data using maximum allowable recycling ratios of 5% and 20% while keeping the maximum allowable recuperation ratio at 5%. In all cases, the measured-to-given dose ratio stays well within 10% of unity, and all overdispersion values overlap at 1 sigma. These results suggest that our rejection criteria have no significant impact on the major conclusions of this study.

10. 384: Measured-to-given dose ratios were 0.99 ± 0.02 , 0.91 ± 0.03 and 0.96 ± 0.03 for Rocks #2, 10 and 18, respectively suggesting that the IR50 SAR protocol is suitable for the Coal Valley limestone samples.

However, what about the pIRIR signal? Additionally, have corrections for residuals been applied to the data? Clarification on this would strengthen confidence in the suitability of the pIRIR protocol for these samples.

We did not use the post-IRIR signal to date the limestone samples because the IR50 signal proved adequate (the signal was suited to SAR, the fading rate was relatively low, and the IR50 signal is less prone to partial bleaching than post-IRIR signals). Because our modern limestone rock samples could not be measured (due to insufficient signal as noted above), we could not determine whether a residual signal remains in well bleached rocks and if a residual dose should therefore be subtracted from the doses measured from our ancient samples. We therefore relied on the SG age distributions (their shape and overdispersions) to inform on the completeness of bleaching of the limestone rock surfaces, as well as the age comparisons between the IR50 ages and the radiocarbon age of the ridge.

11. 593: As expected, IR50 uncorrected ages are significantly younger than pIRIR290 uncorrected ages (Figs 10 and 11) and this is attributed to the high rate of fading of the IR50 signal as well as the lower bleaching rate of the pIRIR290 signal.

However, attributing the age difference to the lower bleaching rate of the pIRIR signal indirectly suggests that the signal may not have been fully reset prior to burial.

We agree that the difference may be attributed, at least in part, to incomplete bleaching (and the age differences between top and bottom sides of the volcanic rocks, Section 4.4.6, also supports incomplete or heterogeneous bleaching). However, the uncorrected ages in Figs 10 and 11 do not allow us to evaluate the effects of bleaching on the IR50 and pIRIR signals because fading has not been corrected for, and existing

fading correction models cannot be applied to such high fading rates as observed in the IR50 data from the volcanic rocks.

12. 792: After ridge formation in this scenario, the pIRIR290 signal that accumulated at the center of the gravels during burial may have been less prone to depletion during subsequent brief periods of sun exposure during bioturbation events, which preferentially depleted the signal near the surface of exposed rock surfaces.

It would be beneficial to present the age-depth plot with a logarithmic scale on the y-axis. This adjustment may reveal that the IR50 signal is also bleached progressively towards the interior, consistent with the bleaching observed in the pIRIR signal. This bleaching may have been further enhanced after burial. Such evidence would further justify the exclusion of surface slices from the dating analysis.

This is a great suggestion. We have replotted all age-depth plots with a logarithmic scale on the y-axis (see Figs 12-14 attached). We also highlight the significance of the IR50 age-depth profile shapes in paragraphs 1 and 2 of Section 4.4.6.

13. Table 4: Could you please include the dose recovery ratios here?

Table 4 DRT ratios have been updated.

14. Figure 12, 13, 14: It is difficult to see the IR50 data. You may consider using a log scale.

A log scale is now included in Figs 12-14.