

Reviewer #2 Specific comments (authors' responses are shown in red after each numbered reviewer's comment):

1. Page 2 line 38, Introduction: "rely on the capacity of". Its not capacity, its property of defects. Change appropriately.

- We will change it to 'property'.

2. Page 2 line 40, 41, Introduction: "In the laboratory, the total amount of energy stored" and "energy absorption rate (dose rate, Gy a-1)". Energy per unit mass is dose.

- We will correct both instances to 'energy per unit mass' and 'energy absorption rate per unit mass'

3. Page 2 line 43, Introduction: "quartz because of its high abundance and resistance to weathering". Besides these two, the fast to bleach OSL signal makes it most appropriate for geological dating.

- We will add this advantage.

4. Page 2 line 45, Introduction: "considering a low dose rate of 1 Gy ka-1". Why to mention 'low' here? Please delete low

- Done.

5. Page 2 line 64, Introduction: "saturation cap at around 1500 Gy, reducing" cap can be deleted

- Done.

6. Page 3 line 80, Method rationale: "study, between one and three grains (~..... curves". Statement not clear, consider revising. What is the reason for bad match? Bad match is often observed due to sensitivity changes. Pls refer (Varma, V., Biswas, R.H., Singhvi, A.K., 2013. Aspects of Infrared Radioluminescence dosimetry in K-feldspar. Geochronometria 40, 266-273.)

- We will change 'between one and three grains' to 'between one and three grains per sample' to clarify and add the suggested reference to explain the bad match.

7. Page 3 line 83, Method rationale: "grains that emitted a detectable signal displayed the expected decay shape." What is meaning of expected shape here? Exponential decay?

- We will change 'expected decay shape' to 'decay shape expected for IR-RF (decreasing signal with increasing dose)'

8. Page 3 line 89, Method rationale: "With dose exposure, the 955 nm emission increases and overlaps with the 880 nm peak." Why is it so? How dose increases the 955 nm emission, does it mean more 955nm recombination centres are being regenerated? Does it reflect multiple trap system?

- There is a mistake in this sentence; it is the ~700 nm emission that increases with dose, not the 955 nm. We apologise for this oversight, we will delete this sentence from the paragraph and replace it with "... which partly overlaps with the 880 nm peak".

9. Page 4 line 101, Method rationale: "The thermal stability of the ~710 nm emission has been, however... the measured IR-RF" The red TL emission in feldspar is generally considered more stable than conventional IRSL method, why the red IR-RF is unstable?

- To the best of our knowledge, no study can be found in the literature specifically investigating the differences of the red-TL and red-RF emissions. Prasad and Jain (2018) described that there are at least two and possibly more components in the red photoluminescence emission, so it's possible that different signals are obtained with TL and RF. This seems to be in line with the findings from Dütsch and Krbetschek (1997) and Krbetschek et al. (2002) who found that the red-RF peak positions shift with potassium concentration. Another considerable experimental difference is that the red RF is obtained at a relatively low temperature of (up to) 70°C (if following the current dating protocol), whereas the red TL peak is above ~200°C (e.g., Zink and Visocekas, 1997).

Prasad, A.K., Jain, M. Dynamics of the deep red Fe³⁺ photoluminescence emission in feldspar. *Journal of Luminescence* 196, 462–469. <https://doi.org/10.1016/j.jlumin.2017.11.051>, 2018.

Dütsch, C. and Krbetschek, M. R.: New methods for a better internal 40K dose rate determination, *Radiation Measurements*, 27, 377–381, [https://doi.org/10.1016/s1350-4487\(96\)00153-9](https://doi.org/10.1016/s1350-4487(96)00153-9), 1997.

Krbetschek, M. R., Götze, J., Irmer, G., Rieser, U., and Trautmann, T.: The red luminescence emission of feldspar and its wavelength dependence on K, Na, Ca – composition, *Mineral. Petrol.*, 76, 167–177, <https://doi.org/10.1007/s007100200039>, 2002.

Zink, A.J.C., Visocekas, R. Datability of sanidine feldspars using the near-infrared TL emission. *Radiation Measurements* 27, 251–261. [https://doi.org/10.1016/S1350-4487\(96\)00141-2](https://doi.org/10.1016/S1350-4487(96)00141-2), 1997.

10. Page 4 line 111, Method rationale: "The μ -XRF and μ -XANES techniques are best suited for this purpose by producing high-resolution maps of elements and their oxidation states". It will be good to provide some details about the mentioned techniques and their usefulness for present work.

- We will add an introduction to the two techniques.

11. Page 3 line 83, Method rationale: "The use of synchrotron μ XRF allows us to improve the spatial resolution compared with previous uses of μ XRF (e.g., Buylaert et al., 2018) by reducing the beam spot size from ~25 μ m to 1 or 0.5 μ m". It is indeed impressive that spot size is smaller and we can work at higher resolutions, but how will it effect S/N ratio and thus elemental concentration estimation? In addition, since spot size is smaller, only few grains analysis may be possible. In such cases, how can we get the statistical representation of entire grain population just based on few grains studies?

- The signal to noise ratio can be optimised by changing the dwell time (signal integration time) for each position without loss of spatial resolution. We did not have issues with the signal to noise ratio for the μ XRF spectra. However, as the reviewer pointed out, the length of measurements at high spatial resolution is indeed a hindrance to obtain a statistically significant representation of a sample. For this study, we chose a high spatial

resolution on a small number of grains to investigate what effects small-scale heterogeneities might have, but this approach might not be appropriate for other studies.

12. Page 4 line 128, Material and instrumentation: Normally Tuff samples are expected to contain Fe rich species. Is this a deliberate choice to look the effect of Fe in the samples as 2 out of 5 are tuff samples?

- It wasn't a deliberate choice to investigate Fe-rich samples. The aberrant IR-RF curves were observed by chance when dating samples from Kenya, particularly from volcanic regions that are often challenging for dating. We chose these samples to investigate the reasons behind the unexpected IR-RF curves and because of the volcanic aspect, Fe was an obvious choice of element for us to investigate further. Other samples (e.g., Gi326, H22550) were chosen as 'reference' samples for comparison.

13. Page 5 line 134, Material and instrumentation: Why only the sample H22550 was etched with HF. Why not same is performed for other samples?

- Sample H22550 was prepared by a different laboratory, which includes HF etching in their pre-treatment of K-feldspar. We chose to include the sample despite the slightly different pre-treatment because this sample was used in an important study that assessed the accuracy of IR-RF dating (Buylaert et al., 2012), and, thus, served as a good comparison sample.

14. Page 5 line 134, Material and instrumentation: "Multi-grain and single-grains ... National Laboratory" How correspondence between IRRF and XRF signals is established?

- Individual grains were first measured on a lexsys research luminescence reader to obtain IR-RF (one grain at the centre of a cup) and then transferred to sticky carbon tape on a polymer slide for μ XRF measurements. The position of the grains of the slide was carefully written down to correspond both signals for each grain. No μ XRF was performed on the multi-grain aliquots. We will rephrase the sentence in question to avoid confusion.

15. Page 5 line 143, Material and instrumentation: "2016). Grains were fixed on a polymer microscope slide....." What are spectral and luminescence characteristics of the base material used?

- For the luminescence measurements, we used the standard Freiberg Instruments stainless steel sample cups; the background level (cup emission + system) is shown in Fig. 2e. The XRF emission depth is dependent on grain density and element, but the signal is expected to be dominated by the top-most volume. For Fe, the XRF signal is expected to characterise only the first 30–40 μ m of the grains, which is relatively small compared to the grain diameter of at least 90 μ m, so we don't expect any contribution from the base material. We will add the information about the penetration depth, which was also asked by the other reviewer.

16. Page 5 line 144, Material and instrumentation: "XRF maps were obtained by scanning across pre-selected regions on the grains 90 x 90 μ m maps," What is the basis of ROIs selection?

- We selected square regions on the grains with low topographic changes (i.e., as flat as possible).

17. Page 5 line 149, Material and instrumentation: “resolution of 0.67 μm was achieved by focusing the beam with” Is it Xray beam focussing or luminescence focussing, please specify.

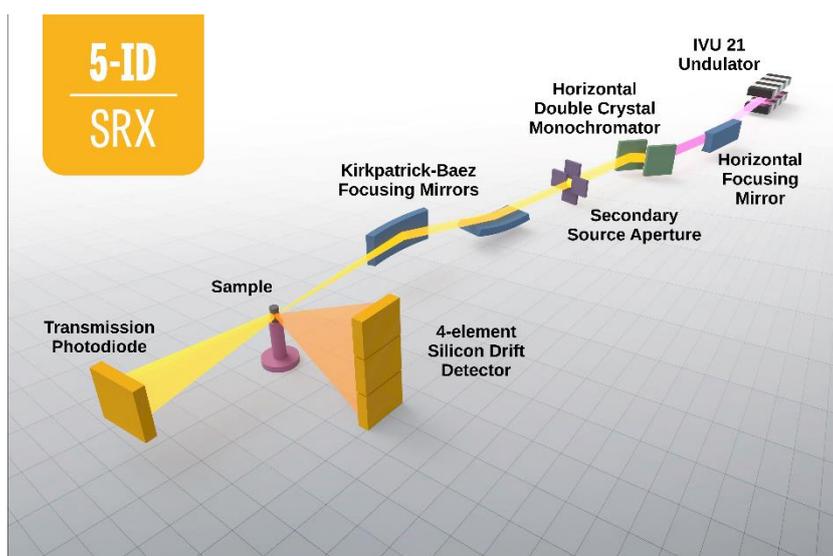
- Only the X-ray beam is focussed onto the sample position to achieve the desired spot size. We will rephrase to “XRF maps were obtained by scanning across pre-selected regions on the grains (90 x 90 μm maps, with step sizes of 0.67 μm and integration time of 0.1 s). The incident X-ray beam was focussed by a pair of Kirkpatrick-Baez mirrors.”

18. Page 5 line 150, Material and instrumentation: “ An incident beam energy of 13.5 keV was” Why this energy chosen any specific reason?

- The incident beam energy determines which elements are ionized, i.e., the beam energy must be higher than the ionisation energies of the elements of interest. There are also elastic and inelastic scatter background peaks that appear near the incident beam energy in the XRF spectrum, so we chose 13.5 keV to provide some separation between the background peaks and the Pb peak that appears at 10–11 keV.

19. Page 5 line 150, Material and instrumentation: “fluorescence was detected through the sum of 4 silicon drift detectors” Why these four detectors were used? Can we use PMT instead? whats the advantage we get with use of these detectors. Can you provide geometry of measurements and experimental setup?

- These detectors form part of the beam line setup and cannot be replaced for XRF measurements. With the silicon drift detectors, we are able to obtain the energy of incoming photons, needed for subsequent analyses. In contrast, PMTs are generally capable of detecting low-level light, so can not be used for detecting X-rays. The experimental setup is shown in the image below and more details on the synchrotron beam line can be found in Chen-Wiegart et al. (2016).



(Image: <https://www.bnl.gov/nsls2/beamlines/beamline.php?r=5-ID>)

Chen-Wiegart, Y.K., Williams, G., Zhao, C., Jiang, H., Li, L., Demkowicz, M., Seita, M., Short, M., Ferry, S., Wada, T., Kato, H., Chou, K.W., Petrash, S., Catalano, J., Yao, Y., Murphy, A., Zumbulyadis, N., Centeno, S.A., Dybowski, C., Thieme, J., 2016. Early science commissioning results of the sub-micron resolution X-ray spectroscopy beamline (SRX) in the field of materials science and engineering. AIP Conference Proceedings 1764, 030004. <https://doi.org/10.1063/1.4961138>

20. Page 5 line 152, Material and instrumentation: “All XRF measurements were normalised to the corresponding incident X-ray flux ” X-ray sources are normally found inhomogenous spatially and temporally. Does this can effect your measurements?

- Spatial inhomogeneity is not an issue due to the submicron beam diameter, but the temporal inhomogeneity needs to be corrected for. That is what the X-ray flux normalisation achieves (example shown in Fig. S2). However, by comparing Fig. S2a and c (before and after correction for flux fluctuation, respectively), we note that the effect is almost imperceptible in our maps.

21. Page 5 line 150, Material and instrumentation: “The XANES maps had a resolution of 0.5 μm (60 x 60 μm).” Are units correct? How does 60 μm X 60 μm translate to 0.5 μm ? not clear.

- The units are correct. These maps cover an area of 60 x 60 μm in steps of 0.5 μm in either direction, thus creating a grid with 120 x 120 data points (stemming from $120^2= 14400$ spectra). We will rephrase the sentence to avoid confusion.

22. Page 5 line 150, Material and instrumentation: “we varied the incident beam energy according to the absorption edge values obtained from the μXANES measurements of Fe standards (Fe foil, pyrite, hematite).” How the specific absorption edge values were estimated?

- The spectra from the μXANES measurements of Fe foil, pyrite and hematite were processed using the software ATHENA v.0.9.26 (Ravel and Newville, 2005) and are shown in Fig. 7. The portion of the spectra containing the absorption edges is highlighted in the inset of Fig. 7. The specific absorption edges (i.e., the excitation energy necessary to eject the electron) were directly read from that figure and are marked with dashed lines for Fe^{2+} and Fe^{3+} . Since the absorption edge of Fe^{3+} is higher than for Fe^{2+} , an XRF measurement at that energy will contain emissions from both Fe^{3+} and Fe^{2+} .

23. Page 5 line 150, Material and instrumentation: “(i) the total Fe (at 7.275 keV),” The energies mentioned here are quite precise. How much is normally the resolution. Since electronic energy levels of specific elements are quite low in energy (~few eVs) compared to what is being provided, then why this much precision is needed?

- The resolution is 1 eV. The energy difference between the Fe^{2+} and Fe^{3+} absorption edges is of 12 eV (equal to 0.012 keV) (see Fig. 7 and lines 155-156), so that precision is needed to differentiate between the two oxidation states.

24. Page 6 line 165, Results: “500 Gy succeeded by an increase, roughly following a saturating exponential shape” What is reason behind increase to saturating exponential behaviour? Why should there be an increase at all considering physics aspect? What is the nature of sample X7343, Is it similar to volcanic tuff?

- Sample X7343 is from a volcanic region, but not directly from a volcanic deposit. Our study could not conclusively define the reason behind the increasing RF behaviour, but we determined that the dual (decreasing then increasing) behaviour of the multi-grain aliquots was caused by superposition of signal from different grains, which in turn were either only increasing or only decreasing. We observed that grains with relatively high Ca and high Fe content tended to display the increasing RF behaviour. It is possible that the increasing signal observed in the IR region is in fact only the tail of the far-red emission associated with Fe, which is known to have this behaviour (see section 2).

25. Page 6 line 169, Results: “contamination, potentially coming from coating around the grains, we” Why is it assumed that coating could be responsible?

- We observed that the grains had a reddish hue, so hypothesised that an iron coating could be affecting the signal.

26. Page 6 line 171, Results: “Despite using density ... high Fe content”. Does that mean it is Na or Ca feldspar grains? Have you performed XRD analysis on bulk to see the mineralogy of samples?

- We have not performed any other analyses on these samples, so cannot say what the bulk composition is. However, our work indicates that a bulk-approach might be insufficient to describe these samples, given their heterogeneous luminescence behaviour at single-grain level.

27. Page 6 line 176, Results: “their visual appearance under a microscope” What were the visual indicators considered for choosing K-Feldspar?

- The visual indicators are given in line 177-178: “Between ten to thirty grains were placed onto two aliquots, one for transparent shiny angular grains and one for white-pinkish rounded grains.” As shown in Fig. 1, the transparent shiny angular grains gave a signal corresponding to that expected from K-rich feldspar.

28. Page 6 line 178, Results: “grains. The regenerated IR-RF signals showed a clear distinction between the two aliquots (Fig. 1), proving it is possible to separate the two observed IR-RF shapes.” This is quite a qualitative way. I am not sure how to progress using only visually inspected grains. The visual appearance and selection can vary depending upon geological settings of grains as well as person observing them. Is there any other rigorous way of making such selection?

- The visual appearance only served as proof of concept and we do not recommend it for future applications. As stated in the conclusions (line 243-246), “manually picking individual grains was necessary to isolate K-feldspar grains, which is unrealistic for routine dating applications in a low-light laboratory. A more realistic way to remove such contamination is by selecting K-rich feldspar grain populations by isolating the emission signal of individual grains with an imaging system.”

29. Page 6 line 184, Results: “presumed to be the low-K, Fe rich minerals identified via SEM-EDS” Low K means possibly high Na or Ca, why only Fe is considered. Fe if present should be in form of defects, which should be in ppm level. Can uXRF measure to such low concentration levels? If Fe is appearing as major element in feldspar separates, it means it is

present in stoichiometric formula and in that case, mineralogy of sample would be different. Please suggest if I am missing something.

- The SEM-EDS Fe-content of sample X343 ranged 0.4–22.4 wt.%. Unfortunately, the μ XRF is not calibrated to provide weight percentages. Both the SEM-EDS and μ XRF analyses were of whole grains, so it is possible that the Fe-rich grains are “normal” feldspar grains (i.e., with no stoichiometric Fe) but with an Fe coating contributing to the signal. It is also possible that these are grains of an entirely different Fe-rich mineral. Further work is required to identify the mineral(s) with aberrant IR-RF behaviour.

30. Page 6 line 186, SG IR-RF characterisation: “signal of twenty-two individual grains coming” Can you specify mineralogy of each grains, which are picked for such measurements?

- These were “new”, previously unmeasured grains, so we do not have SEM-EDS analyses for them.

31. Page 6 line 165, SG IR-RF characterisation: “2): Category #1 for grains with a decreasing IR-RF signal, category #2 for grains with an increasing IR190 RF signal, and category #3 for grains with a flat signal”. How many grains falls in each category and is there any link to the provenance.

- The numbers of grains are listed in Table 1 (9, 8 and 5 grains, respectively, in each category). The Kenyan samples had a higher proportion of category #2 grains, but the numbers are too small to make any meaningful interpretations of provenance.

32. Page 6 line 165, SG IR-RF characterisation: “we also observed the unwanted decreasing IR-RF signal for one of the four grains for sample H22550, which is from a coastal sand deposit.” We expect IR-RF signal to decrease with irradiation, so why it is said unwanted ?

- We apologise, we meant “the unwanted increasing...”

33. Page 6 line 193, SG IR-RF characterisation: “When the total signal of the theoretical aliquot was composed of more than 50% of signal from the category #2 grain, we observed the same decay shape as in figure 1 for a multi-grain aliquot sample X7343” Obviously, since the two different category of grains having two different IR-RF characteristics are being added, so the result will depend on the proportion of the individual populations in the mixture. More importantly, it is important to know, how these two grains are different with respect to crystallography or stoichiometry or defects concentration. Is the nature of curve repeatable over repeated bleaching and irradiation cycles?

- Future work should indeed investigate in detail the mineralogy, crystallography and trace elements in the grains leading to the two signals. However, we do not agree that the mixture of decay shapes is trivial. We modelled that the IR-RF signal from a multi-grain aliquot with up to 20% contaminant grains would not be flagged as abnormal (i.e., it would still display only a decreasing signal), but the IR-RF characteristics relevant for dating (e.g., saturation, stability) would already be impacted. This finding might explain the mixed accuracy observed for IR-RF ages. The curve shape is repeatable.

34. Page 7 line 201, SG IR-RF characterisation: As mentioned by authors, long-term signal stability may not be there for bad traps, is there a way to prove it? How do we know it without experiment?

- We have not yet conducted IR-RF fading tests for these samples, but merely mentioned what the consequences would be of different luminescence behaviours between the two groups.

35. Page 7 line 204, SG IR-RF characterisation: “Further, our results demonstrate...” I agree with this statement, but it is still not clear how can we segregate K-Feldspar and other minerals. Manually it will not be possible on routine basis.

- Using an imaging system, as suggested in the conclusion (lines 244-246), would solve this issue. Dose-response curves can be obtained for all grains and only those with appropriate characteristics would be used for D_e estimation.

36. Page 7 line 207, Subgrain μ XRF elemental maps: “We then fitted each of the **18 225** spectra for....” this statement is not clear

- We meant eighteen thousand two hundred and twenty five spectra: one spectrum for each pixel in a grid of 135 by 135 pixels with step sizes of 0.67 μ m, covering an area of 90 by 90 μ m. Each of these spectra was fitted to deconvolute the emission peaks of different elements.

37. Page 7 line 212, Subgrain μ XRF elemental maps: “characterise visible inclusions (see Table 1).” The number of grains analysed per samples are quite small to represent the statistics of system. Can we consider them as representative of whole samples? It is difficult to conclude unless sufficient data points exists.

- It is true that we had only a small number of grains per sample. However, joining all samples and categorizing by IR-RF behaviour gave us 9 and 8 grains in groups #1 and #2, respectively, which we believe to be sufficient for this pilot study. We observed the trend that grains with the desired IR-RF signal shape had high Pb and Ba and low Fe and Ca contents. In future work, we will increase the number of grains.

38. Page 7 line 214, Subgrain μ XRF elemental maps: “all contain K, Pb, Fe and Ba, among other elements (Fig. 4).” What is typical concentration of these elements? Considering K is a major element present in stoichiometric formula, how much is relative concentration of the other elements?

- Unfortunately, the system is not calibrated to convert the XRF intensity to elemental composition (e.g., in weight) of the different elements. We can only compare the XRF intensities of each element, between different grains, as done in Fig. 5.

39. Page 7 line 223, Subgrain μ XRF elemental maps: “grains from category #3 cluster relatively close to those from category #2, suggesting” How and why does this effect the IR-RF properties? These are observations, but what is the reason for IR-RF signal due to such clusters is not clear.

- We do not have a definitive explanation for the different IR-RF behaviours in grains of similar elemental compositions. We suggested that “elemental composition alone is not responsible for the lack of IR-RF signal but that the spatial configuration of the elements is another important factor to consider” (line 223). By this, we mean that the presence of the same elements at different sites in the crystal might elicit different luminescence

responses. Site-selective luminescence behaviour has been reported previously (e.g., Kumar et al., 2020).

40. Page 7 line 237, Mapping oxidation states with μ XANES: “suggest that Fe exists on the surface of this feldspar grain in its Fe³⁺ and Fe²⁺ states.” It is great that using uXANES, we could map the presence of Fe on the surface of feldspar grain, but luminescence or IR-RF is a volumetric phenomenon. How this observation is helpful in explaining the IR-RF signal.

- In future work, we aim to set up a dual detector to detect simultaneously XRF and the IR-RF resulting from the X-ray irradiation. Such a setup would enable a direct comparison.