

## Reply to Reviewer 1

We thank Reviewer 1 for his/her constructive and helpful comments, allowing us to improve our manuscript. Below, we provide a detailed response to the raised concerns and suggestions. For clarity, quotes from the review are italicized.

Specific comments:

- *Because of the algorithm used in  $\mu$ Dose-system, the U and Th contents are negatively correlated. We can see from Figures 3, 4, 7, 8 that, when the Th content or activity is higher (assume it is overestimated), the corresponded U is lower (underestimated), and vice versa. I think it might be more helpful to add the bulk U+Th activity as another parameter for comparison. Even though the individual activities of Th and U are deviated from the expected values, as long as the bulk U+Th activity is close to the expected value, it might still be treated as a successful measurement regarding the calculation of environmental dose rate. The conversion factor from activity to beta dose rate is higher for U and lower for Th, while the conversion factor from activity to gamma dose rate is higher for Th and lower for U. They compensate each other, and the total dose rate does not vary much with the exact Th/U ratio (e.g. section 4.3.1 in Aitken 1985; Li and Tso, 1995). For example, in thick source alpha counting (TSAC), sometimes we can simply assume that the sample has equal activities of U and Th series. So, I think the bulk U+Th activity would be another evidence for the reliability of  $\mu$ Dose-system to accurately determine the environmental dose rate.*

We thank Reviewer 1 for this interesting comment and added plots that show bulk U+Th activities of the investigated loess standards Nussy and Volkegem. These additional plots were combined with plots showing simulated dose rates calculated for the standards. Furthermore, the results shown in these plots are discussed in the manuscript.

- *Maybe, it would be even more straightforward to calculate the final environmental dose rates for comparison. Dose rates can be calculated according to the true settings of individual samples (grain size, mineral, water contents, etc). Alternatively, dose rates of all samples can be simply calculated based on etched quartz with a fixed diameter (e.g. 150-200  $\mu$ m) by U, Th, K measured from  $\mu$ Dose and other methods (TSAC, ICP-OES, gamma spectrometry), and just assume constant cosmic ray and water content. Comparison of dose rates can directly give the readers an impression about the performance of  $\mu$ Dose in determining the environmental dose rate.*

Done. We calculated environmental dose rates for all analysed samples including loess standards as well as natural samples. For the sake of simplicity, we followed the suggestion proposed by Reviewer 1 to use assumed constant values for cosmic radiation ( $0.150 \pm 0.015$  Gy/ka) and water content ( $15 \pm 5\%$ ). Since these calculated dose rate values do not represent the actual environmental dose rates that might have been derived for the various locations when applying the actual water contents and cosmic radiations, the calculated dose rates are called ‘simulated environmental dose rates’ in the manuscript and in the plots. The details of dose rate calculation are described in the figure captions and the results are discussed in the manuscript.

*Aitken, M.J., 1985. Thermoluminescence dating. London. Academic Press.*

*Li, S.H., Tso, W.M., 1995. Systematic error from Th/U ratio in luminescence and ESR dating. Nuclear Science and Techniques 6, 113–116.*

- *To study the impact of measurement duration on the results, the authors have made repeated measurements with different durations. As the data can be stored during the measurements, I have a concern: are the short-measurements separate measurements or are the short measurements the former parts of long-measurements? I guess the former strategy is more reasonable, otherwise there would be correlation between the short and long-measurements.*

We agree with the reviewer that there would be a potentially problematic correlation between short-, medium- and long-time results if the short- and medium-time results were derived from only one long-lasting measurement. Therefore, all analysed measurements were performed as separate measurements. We did not use a long-lasting ‘master’-measurement to derive different short-term parts. In order to clarify this aspect, we added an explanation to the manuscript.

- *When comparing the results of the 47 natural sediment samples, the discrepancy of several samples is attributed to the disequilibrium in U and Th decay chains. For example, fluvial flood plain sediments may have strongly alternating ground water levels which can increase or/and decrease specific radioactive daughter nuclides in the U and Th decay chains (line 483). In the beginning, I thought you meant the radon-loss induced disequilibrium, then I got confused. Because for the TSAC and gamma spectrometry measurements, the samples have been stored for 4 weeks before measurements and for  $\mu$ Dose-system the samples have not been stored before measurements. If the discrepancy is caused by radon, the problem would exist for all samples. Now, I guess you meant the disequilibrium caused by long-lifetime daughter nuclides, right? Could you give examples of the daughter nuclides that might be influenced by the ground water level change, and if possible, list a reference? That may help the readers to better understand what you mean.*

The reviewer is right when assuming that we meant disequilibria caused by long-lifetime daughter nuclides. We added some sentences to the manuscript clarifying this point and illustrating the complex character of such potential disequilibria. We also added some references.

- *In the sample preparation step (line 195), the samples were pulverized in a ball mill (29.5 Hz for 45 minutes) and then dry sieved restricting the grain size diameter to < 63  $\mu$ m. Do you assume that grain size < 63  $\mu$ m would be fine enough for alpha counting? And usually how much sample would be left coarser than 63  $\mu$ m after being pulverized in a ball mill for 45 minutes? I am a little worried that this sieving step may cause fractionation of the sample component. For example, if the quartz is more difficult to grind than feldspar, the left residue of > 63  $\mu$ m will contain more quartz and the fine powder will have higher K (as well as Th, U) contents. Or maybe, the left residue of > 63  $\mu$ m contains more heavy minerals which have high Th, U contents (e.g. zircon), and the fine powder will have lower U, Th contents. Would it be better that we extend the grinding time and avoid the sieving process?*

We would like to thank the reviewer for this helpful comment pointing to a misleading verbalisation in the manuscript. When Reviewer 1 is asking “[...] how much sample would be left coarser than 63  $\mu$ m after being pulverized in a ball mill for 45 minutes [...]”, I think he/she is most probably implying that a prolonged milling of 45 minutes will provide widely pulverized materials. And in fact, the reviewer is right in that assumption. The sieving step has been introduced in the Giessen Luminescence Laboratory as additional backstop for the sample preparation. If the amount of coarse-grained material is rather large (i.e., if there is any residual

> 63 µm) this indicates that the applied milling duration was not sufficient and that the whole sample (including the particles > 63 µm) should be subjected to further milling. Thus, the sieving step is not primarily intended to exclude grains with diameters > 63 µm, but to serve as an additional step to survey the quality of our preparation procedure. We added some additional sentences to the manuscript to clarify this point.

Up to now, residuals > 63 µm have been negligible for the utmost number of samples investigated in our laboratory when applying our long-lasting milling duration. In order to check this for the samples investigated in the submitted study I enquired our laboratory logbook in which our laboratory assistant records input (total sample) and output (< 63 µm) masses for all samples. Overall, only few samples so far investigated in Giessen showed suspicious results suggesting that the prolonged milling-procedure was not sufficient to pulverize the sample. None of these samples were used for the submitted study. As a result, we are confident that the impact of the sieving procedure can be regarded as negligible for the samples investigated in the submitted manuscript. This aspect was also amended to the manuscript.

Technical corrections:

*Line66: 'disk' is used here while 'disc' is used in line 197. Please make it consistent.*

Done. Everything was changed to 'disc'.

*Line79: 'These pairs are the result of...' change 'result' to 'results'.*

Done.

*Line174: '16.5 ± 1.5 mg/kg for K', change to 1.65 ± 0.15.*

Done. We changed the value to 1.65 ± 0.15 as well as the unit to % to make it consistent with the values used in the tables.

*Line 197: 'sample carrier', could you please indicate in Fig. 1a which one is sample carrier? And also give the name for that metal base.*

Done.

*Line 385: 'of only view hours', change 'view' to 'a few'.*

Done.

*Line 443: I think it is not necessary to use a separate paragraph here.*

Done.

*Table 4: Maybe it is better to also convert the activities of Murray et al. (2018) into the concentrations. That would give the readers a direct comparison of the results measured by different methods or labs.*

Done. We converted the activities provided by Murray et al. (2018) to concentrations. Additionally, all concentration values for which no activities were provided in the original publications were converted to Bq/kg. We used conversion factors provided by Guerrin et al (2011). All converted values are highlighted using a \*\*-symbol which is explained in the table annotation. Additionally, we added a column showing ‘simulated environmental dose rates’ for each sample. Since IAG did not provide any information on the potassium content of the investigated Nussy standard, no dose rate was calculated for this dataset. The dose rates were calculated with DRAC (Durcan et al. 2015) based on the originally provided values for U, Th and K and assuming constant values for cosmic dose rates ( $0.150 \pm 0.015$  Gy/ka) and water contents ( $15\% \pm 5\%$ ).