

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

As mentioned in my first review the study “Investigation of quartz ESR residual signals in the last glacial and early Holocene fluvial deposits from the Lower Rhine” is interesting and worth being published. I congratulate the authors for their work. Furthermore, the manuscript was significantly improved following the reviewing process.

We thank the reviewer for the kind words on the improvement of the manuscript.

I have only one comment left. I am uneasy with the term “sensitivity change”. I find the term a “black box”. While in luminescence we have at least models of what might be found in that back box, in ESR, in my personal view, there are no models put forward yet. As such, I find it luminescence jargon transposed without a clear basis yet. By this comment it is not my intention to impose to the authors not to use the term, but as an intro to the dose recovery test interpretation that I am still not fully satisfied with. In the introduction the authors state that “Previously, ESR dose recovery tests have been conducted by Beerten et al. (2008) on quartz derived from dune sands and Asagoe et al. (2011), who used quartz from tephra samples. Unfortunately, both studies use an intensive thermal treatment (annealing) of the sample to erase the natural signal before artificial irradiation, which reduces the significance of the test.” However, here a preheat is applied as well. Moreover, an unknown value (the D_e) is fed into the protocol as a known one, as explained in my previous comment. While in principle I agree with the preheat, I am still not convinced that the way this dose recovery test was conducted is the best possible one. Maybe it would have been better to subject the sample to extensive bleaching experiments until (hopefully) a constant residual would be reached, and then add the 1000 Gy dose. I understand this is extremely time consuming and I do not impose it now either. However, I think a caveat in this regard should be added in the text. In my view, while dose recovery tests should be incorporated into ESR dating, there is still work to be done before an exact protocol is established.

To avoid any misunderstandings, we have deleted most of the phrase “sensitivity change” from the manuscript and explained in other words.

I do not fully grasp what is meant by "However, the result of the dose recovery test suggests that the ratio of bleachable/unbleachable components should be compared before and after the annealing step, in order to understand the problem of the dose recovery test." in the discussion. It might be better to move this information to the dose recovery section and expand it for clarity.

We have added a more detailed explanation in the discussion section of the paper.

Other small corrections:

-paragraph at line 75 subscripts and superscripts are missing: “[AlO₄]₀ [TiO₄/M⁺]₀”

The missing subscripts and superscripts have been added.

- paragraph at line 135 “thermal treatment included in the SAR protocol may change sensitivity of the ESR centres per unit dose” delete per unit dose as sensitivity itself is signal per unit dose.

We rephrased the sentence without using “sensitivity”.

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

The manuscript focuses on residual ESR signals in quartz (Al and Ti centres) of fluvial deposits from the lower Rhine. In this study, the authors compare newly obtained ESR results with OSL ages (Lauer et al., 2011). In general, the manuscript is well written and has improved based on the reviewer comments. I think this manuscript is definitely worth to publish as it gives valuable information in ESR signals in a fluvial environment, which is important for other ESR dating studies. However, I think the authors can extend the discussion and focus a bit more on the ESR signal properties rather than only presenting residual doses, which might not be fully reliable (poor DRT, sensitivity changes). My main concern focuses on the poor dose recovery test.

We thank the reviewer for the very insightful comments to our manuscript.

1) The dose recovery test is an issue in my opinion. How can we trust a measurement protocol that is proven inaccurate based on the poor DRT?
a) What are the fitting uncertainties of the different centres?

In terms of the fitting uncertainties for the DRT, the R^2 for the Al centre is 0.993 ± 0.002 , for the Ti-Li centre is 0.995 ± 0.001 , for the Ti-mix centre is 0.999 ± 0.001 and 0.994 ± 0.006 for Ti-H centre.

I wonder whether the poor DRT results from the Ti-H centre come from a poor measurement (e.g. low ESR intensity). I appreciate that the authors provided additional information on the measurement uncertainties. I think this is an important point. The Ti-H centre shows an uncertainty of up to 16% and just between the three rotations. It was noted that the ESR intensity of the Ti-H centre was “relatively small”.

I think, in terms of Ti-H and the large uncertainty is it a mixture of the weak natural signal, which results in a low S/N ratio, an angular dependency of the signal itself, and for sure, to a certain portion, human error.

The authors should discuss the reliability of the different ESR signals regarding signal properties and experimental errors. At the moment, the manuscript does not provide any information on the reliability of the signals, except of poor DRT results (Al and Ti-H). Please provide more information on ESR intensities, fitting uncertainties, etc. in the manuscript or SI.

For sample RB-II which was used for DRT, the Al centre signal intensity (a.u.) is ~8 times bigger than Ti-Li, ~7 times bigger than Ti-mix and ~22 times the size of Ti-H.

In terms of the fitting uncertainties for the DRC's (DRT), the R^2 for the Al centre is 0.993 ± 0.002 , for the Ti-Li centre is 0.995 ± 0.001 , for the Ti-mix centre is 0.999 ± 0.001 and 0.994 ± 0.006 for Ti-H centre. The relative standard deviation (1-sigma, %) of the three measurements for sample RB-II natural signal were for Al 0.55, Al after subtraction of peroxy 0.86, Ti-Li 2.27, Ti-mix 2.9 and Ti-H 16.1.

b) What is the saturation level of the Ti-H centre? The added ca. 1000 Gy on top of the natural (residual) signal is in my opinion way too much for the Ti-H centre, which saturates usually much earlier than the Al and Ti-Li centres and often <1000 Gy. Can the DRT work if the Ti-H is saturated? It would be worth showing the D_0 values of the signals.

In terms of the sample used for the DRT, the mean D_0 for Ti-H is 668 Gy. So it is close to $2x D_0$ after adding 963 Gy, but still below $2xD_0$.

2) How do sensitivity changes affect the results? The authors mentioned “Only the significant issue, which can affect the accuracy of the dose measurement is the potential sensitivity change by the annealing step, which was tested by the dose recovery test.” and the “reported residual doses may be overestimated [Al centre]” due to sensitivity changes. In line 105, an annealing treatment of 420 °C for 2 min was chosen based on the results of Tsukamoto et al. (2015), which did not show sensitivity changes. In contrast, sensitivity changes were observed in the current study. To which extent are the Al dose results overestimated due to sensitivity change (75%?)? It should be discussed better in regard to the final residual dose results as a recommendation is given in the conclusion to subtract such high doses.

If the SARA slope is 1.75, yes you can correct the D_e by dividing the D_e with the slope; this is the definition of SARA. But since we only use one added dose point in our dose recovery experiment, we do not want to do this. At the moment, even with the Ti centre (better bleachable signal than the Al centre) we observed significant residual dose – all we can say is using the Al centre to date fluvial sediments is even more difficult.

3) Inter-aliquot scatter is one point, but technical uncertainties (e.g., differences in spectrometer sensitivity, HS cavity) might lead to variations in D_e repeatability. This has often been observed in other ESR dating studies using even larger aliquot sizes (e.g. 200 mg). I know that ESR measurements are very time consuming, but I think it would be worth exploring the variability in D_e values of different aliquots for a single sample. I agree that in the present study the focus lies on the determination of the residual dose, and not on a precise burial date, but given the fact that the residual dose should be subtracted from the sample’s D_e to receive a more precise age, I think it is important to evaluate also the residual dose in a more robust way. Maybe the authors could give a recommendation for ESR measurements in view of future single aliquot measurements. For example how many aliquots should be measured in order to provide a better statistical estimate of the D_e (otherwise it will be just one D_e value that might not be correct).

We do not really think it is possible to obtain a precise residual dose from modern analogue samples. Our result from different samples using one aliquot each from the same terrace is similar enough to give a ball park residual dose from the Rhine sediments. As we did not measure multiple single aliquots for the samples, it is not possible to make such recommendation in this paper.

4) Although the authors have included Table 1 showing a brief sample description, the manuscript still does not show any discussion on sediment process properties of the samples in relation to their bleaching kinetics. As sediment properties seem to be different (Table 1) and residual dose variations have been observed (e.g., Ti-Li ~400-900 Gy), it would be interesting if the authors could comment on this in the discussion, highlighting bleaching differences of different fluvial transport modes. It is important which kind of fluvial deposit was measured as the transport mode has a large impact on the bleaching of the signals. More information should be given in the discussion.

We have added a paragraph on the samples' deposition environments and their bleaching kinetics with respect to the ESR residual doses we measured to the discussion part of the MS. We did not find any dependency of the ESR residual with the deposition environments.

5) Line 148: I agree that the subtraction of a total residual might give us an idea about the bleached ESR signal. However, it remains difficult to find a representative modern analogue for Pleistocene samples, and if there is no appropriate sample available we still deal with two unknowns: the partially bleached component and the optically unbleachable component. So I think it is difficult to get a reliable residual dose that can be subtracted from older samples, and measuring an artificially bleached aliquot to estimate at least the unbleachable Al component is important. "residual subtraction is highly recommended" (line 148) should be rephrased, as this is only valid when a representative modern analogue sample is available.

The sentence has been rephrased according to the reviewer's suggestion: "The results clearly show that the ESR D_e for all samples are significantly larger than the OSL D_e of Lauer et al. (2011) and therefore residual subtraction is highly recommended if a representative modern analogue sample is available."

6) Lines 148-150: „the observed residual doses confirm the trend in the signal's bleaching behaviour“. This is only true when the results of all centres are reliable, but as mentioned before the Al is probably overestimated and the Ti-H centre results are probably underestimated. Moreover, the total residual dose of the Al centre contains two components (optically unbleachable component and remnant dose due to partial optical bleaching), so the Al centre is only hardly comparable to the Ti centres as the latter does not contain an unbleachable signal component. Secondly, the Al centre is characterised by slower bleaching kinetics.

We have rephrased the sentence to "Furthermore, the observed residual doses follow the trend in the signal's bleaching behaviour as described by Toyoda et al. (2000): the Al centre shows the largest residual followed by the Ti-Li and Ti-H with the lowest. The size of the residual dose for the Ti-mix lies in between the Ti-Li and Ti-H."

The rough estimate of the size of the unbleachable Al component (line 160) is 90 ± 210 Gy? How is this in agreement with estimations of 500-700 Gy by Tsukamoto et al. (2018) and Timar-Gabor et al. (2020) (line 161)? It remains rather difficult to estimate the unbleachable component without a proper bleaching test. Please clarify that in the discussion.

For clarification, the aforementioned sentence has been extended by "[...], although it is much smaller than the values reported by Tsukamoto et al. (2018) and Timar-Gabor et al. (2020) from aeolian sediments."