<u>Author's response to the review by Jakob Wallinga, Wageningen University / Netherlands Centre for Luminescence dating</u>

Review of Biswas et al., submitted for publication in GChron

Manuscript 'Novel insights into the post-IR IRSL200 signal bleachability of singlegrain K-feldspars in fluvial modern analogues from the Southern Central Andes, Chile'

Review by Jakob Wallinga, Wageningen University / Netherlands Centre for Luminescence dating.

The authors investigate the resetting of feldspar pIRIR200 signals upon light exposure, using data from both modern fluvial samples and laboratory irradiated and bleached samples. They find that bleaching rates vary between grains, and explore whether these differences can be explained from the chemical composition of the grains. The work and insights gained are original and relevant to the luminescence dating and tracing community, and fit well with the scope of the GChron journal. I would therefor support publication of the work. However, I would suggest that major revisions are needed to improve quality and clarity prior to publication. Below I list the most crucial improvements to be made.

We are grateful to Jakob Wallinga for the careful and thoughtful review of our manuscript. The constructive feedback provided has been invaluable in improving our work. In the following sections, we respond to each comment individually and outline the revisions incorporated into the manuscript.

My first concern is the focus of the paper. The manuscript documents a wide range of experiments conducted on a set of samples – a more coherent storyline and choices on which data to include in the main paper would contribute to a more powerful manuscript. At the end of the introduction five aims are listed, which is a lot for a single paper. I would recommend to reduce to a maximum of three aims. In this context, I would suggest to focus on bleachability of individual grains and how it relates to geochemical composition of the grains – as this is the original and unique aspect of the paper. For remnant doses in modern natural samples, the light exposure history (rather than bleachability) is likely to be the decisive factor (as is also discussed in e.g. the associated thesis by Karman—Besson). The remnant dose data is hardly discussed and distracts from the original and important work to investigate bleachability at single grain level. Moreover, I have some serious reservations on the interpretation of the data. As other approaches/assumptions may change the interpretation and hence the conclusions drawn, I would like to re-assess the paper after the authors have implemented the suggestions below – or provide convincing argumentation why they decide not to.

Reply: We have now refined the articulation of our research objectives accordingly. In response to your concerns regarding certain methodological approaches and interpretations, we provide detailed point-by-point clarifications below.

Below I list my major reservations with respect to the data interpretation:

- MAM and choice of (absolute) sigmab: sigmab in the MAM model should account for overdispersion in a De dataset that is not related to heterogeneous resetting of the luminescence signal and to measurement uncertainties. Usually, the main cause for such additional overdispersion is expected to be heterogeneous beta dose rate. When approaching zero dose, as would be expected for the best-bleached grains from modern deposits, overdispersion due to heterogeneous dose rate would be zero by definition. Any remaining signal in partially bleached grains may have accumulated under different dose rates, but well-bleached grains (0 dose) are not affected by this. I therefor strongly object to the assumed 8.58 Gy absolute overdispersion. Moreover, the trick to fit a trendline of absolute OD as a function of De (Fig. S5) could work if all the samples were of different age and well bleached – then absolute OD is expected to be a % of De, and a fit would be expected to go through the origin. Here, both the CAM doses and OD values are artefacts of heterogeneous bleaching. (In spite of that, fitting a fit forced through the origin for S5 would also provide a pretty good R^2). In conclusion: selecting a suitable (absolute) sigmab for very young samples is cumbersome, but for modern samples (as studied here) it should simply be zero. The overestimation of sigmab likely results in a (gross) overestimation of the MAM De; running the MAM with sigmab of 0 Gy would result in estimates much closer to the true burial dose of 0.

Reply: We acknowledge this concern raised by the reviewer. We agree that the use of 8.58 Gy as sigma-b input in the unlogged Minimum Age Model (MAM) is not ideally suited for our modern fluvial samples and is presumably overestimated. The main issue with its application here is that the absolute sigma-b value (which must be specified a priori for the unlogged MAM) is particularly difficult to constrain for young samples with strongly skewed dose distributions resulting from highly heterogeneous bleaching, which also include negative D_e values. After testing different setups, including, for example, the intrinsic sigma-b derived from the dose recovery experiment, we had to come to the unfortunate conclusion that the unlogged MAM is not suitable for our modern analogue single-grain dose distributions.

We have therefore replaced the MAM with an approach better suited to the structure of our data. The mode of the single-grain dose distribution (representing the global maximum of the distribution) is now used as our dose estimate to represent the dominant well-bleached grain population. Unlike the MAM, this requires no assumption about sigma-b. Uncertainty in the modal estimate was quantified using a nonparametric bootstrap procedure (n = 1000). For each iteration, we resampled the dataset with replacement, recalculated the mode, and then derived the one-sigma (68%) confidence interval from the resulting distribution of bootstrap mode estimates. We consider this a more robust and statistically transparent representation of the dominant well-bleached population in our heterogeneously bleached modern fluvial samples. The results are plotted in the revised Fig. 6 of the main text, and indeed some of the estimates are much closer to a burial dose of 0 Gy than previously estimated using unlogged MAM. We changed the corresponding interpretation and discussion accordingly.

Selecting 'representative' grains for bleaching investigation (Fig. 3): here the decay after 1 minute of SOL2 bleaching was used to select the 'representative' fastest (grain 1), slowest (grain 3) and typical (grain 2) bleaching grains. However, grain 1 and 3 should not be regarded as representative, but rather outliers. I expect that the normalised Lx/Tx will have considerable uncertainty (not shown in Fig. 3) and that some of the trends (e.g. bleaching curves in Fig. 3b crossing over) are artefacts. I would suggest to show data for all grains in Fig. 3b (e.g. in light grey), and show trendlines for averages of subsets of grains e.g. the upper quartile, lower quartile, and interquartile for the 1-minute bleach. If the authors decide to show single grains, uncertainties should at least be shown in the graph.

Reply: Thank you for suggesting alternative options. We now show data for all grains with trendlines for averages of subsets based on the upper quartile, interquartile, and lower quartile for the 1-minute bleaching step, which is tracked consistently across all subsequent bleaching steps in Fig 3b and Fig. S8.

- The inter quartile range (IQR; Fig. 4e) is shown to decrease with bleaching time. However, this absolute IQR (Lx/Tx) is not a meaningful value, as Lx/Tx will decrease with bleaching time. Hence the relative IQR (IQR (Lx/Tx) / (Lx/Tx) would be much more meaningful, or alternatively the overdispersion (as %) after different bleaching times.

Reply: Thank you for pointing this out. Following your suggestion, which is in line with the second reviewer, we have revised Fig. 4e to show the IQR based on normalised L_x/T_x values to enable more meaningful comparisons.

- In Fig. 6 an exponential trendline is fitted, suggesting that residual dose (after 2 days of SOL2 bleaching) is related to the remnant dose (MAM-derived) in modern samples. From the fit, the unbleachable component is estimated. However, the data could also be fitted (quite well) with a linear trendline, which would result in a very different (near zero) intercept. Is there any (physical) reason to adopt the exponential fit (of this order)?

Reply: We agree that our original fitting approach required reconsideration. We initially used an exponential fit because it maximised the R^2 value and best described the relationship between MAM D_e and mean residual dose. With our revised approach, using the mode for both D_e and residual dose, a linear model now captures the updated relationship just as effectively. We have therefore applied a linear fit and updated the discussion in Sections 4.1.4 and 4.3.1.

Additional minor comments:

- Writing & storyline: The writing could be clearer and more succinct in places. Some of the information appears in the wrong sections (e.g. line 312-318 should appear in methods rather than results).

Reply: The manuscript has been carefully reviewed and restructured accordingly. Lines 312–318 have been moved from the results section to the methods section (Section 2.5) to ensure proper placement of methodological content.

- Dose recovery: it remains unclear why the measurement temperature is based on a dose recovery test on single-aliquots. The dose recovery in the single-grain measurements is not very convincing.

Reply: We acknowledge the reviewer's concern. The measurement temperature was selected based on single-aliquot dose recovery tests, as these provided the most practical means of evaluating the suitability of the post-IR IRSL protocol across different temperatures prior to conducting the more resource-intensive single-grain measurements. While we recognise that the single-grain dose recovery results are less convincing at face value, we already noted in the main text that when 1-sigma uncertainties are considered, the majority of samples (with the exception of CHLEA-11) fall within the acceptable range of 0.9 to 1.1, suggesting adequate protocol performance for most samples.

- Discussion, section 4.1 – this section is unclear as the reader is not reminded what the starting dose was for the mean residual dose experiment. It is confusing to refer to 'this variability' for the single-grains.

Reply: The variability mentioned in this statement is not derived from the bleaching experiment itself, but rather from residual dose estimates measured after two days of Sol2 bleaching. Our intention was *to demonstrate to the reader that variability in residual dose estimates is evident both across and within samples. To make it reader-friendly, we have rewritten the sentence as "Similar variability is also apparent at the single-grain level, where residual doses (measured after two days of Sol2 bleaching) within individual samples range from near zero to about 23 Gy, as illustrated in Figure 5a for the four samples included in the bleaching experiment".*

- Make sure to differentiate between dose and signal (e.g. line 71-72)

Reply: We have now carefully reviewed the manuscript to ensure clear differentiation between 'dose' and 'signal' throughout.

Author's response to review of Anonymous Referee #2

This study investigates the bleaching behaviour of the post-IR IRSL200 signal in K-feldspar grains from modern fluvial sediments in the Southern Central Andes of Chile. The authors conducted controlled laboratory bleaching experiments at the single-grain level on multiple samples, assessing the influence of bleaching duration, natural dose, geochemical composition, and catchment lithology on residual and remnant doses. They observed substantial variability in residual doses both among different samples and among individual grains within the same sample, but found no clear control from geochemistry or catchment lithology. The study also highlights limitations in correcting palaeodoses using remnant doses from modern analogues and evaluates different approaches for palaeodose correction. Overall, the work has important implications for improving age accuracy, especially for young samples in dynamic depositional settings. The methodological details are clearly presented, the manuscript is well written, and it fits well within the scope of the journal.

We thank anonymous referee #2 for reviewing our manuscript and for the constructive suggestions that have helped improve the quality of the work. Below, we provide point-by-point responses to each comment.

However, my main concern is the choice of the sigma-b value used in the minimum age model. The σb value (8.58 Gy) is obtained by extrapolating a linear relationship between absolute overdispersion and CAM De, but this approach needs more explanation. The reported R^2 of 0.58 (Fig. S5) shows only a moderate correlation, meaning that much of the variability in overdispersion is not explained by dose. Under these conditions, using the y-intercept as a representative overdispersion for a well-bleached grain population is not very convincing. This approach would also only work if all samples were well bleached and of different ages. However, the Kernal density of single grain De values showed that these samples (at least for four samples that showed in Fig. S6) are not well bleached. The authors should therefore clarify the reasoning behind this approach.

As an alternative, the expected overdispersion for well-bleached grains could be estimated by separating the sources of variability into internal and external (presumably dominant by those caused by beta microdosimetry) components. Then one can estimate the expected overdispersion by combining these two components in quadratic. The internal variation is easy to obtain experimentally (for example by using the average overdispersion from the dose recovery tests of the four samples). For well-bleached modern samples, the external microdosimetry can be assumed to be negligible, as these grains should theoretically have near-zero doses.

Reply: We acknowledge the reviewer's concern regarding the use of 8.58 Gy as sigma-b input in our unlogged Minimum Age Model (MAM). We agree that selecting an absolute sigma-b value for our modern samples is inherently difficult, as also suggested by reviewer-1. Although the traditional MAM is widely applied to feldspar, we believe it is not well-suited to our single-grain post-IR $IRSL_{200}$ feldspar dose data of modern fluvial samples, which show highly heterogeneous bleaching and strongly skewed dose distributions, and at the same time, a significant portion of grains with negative dose values. After exploring multiple configurations, including the use of intrinsic sigma-b values derived from a dose-recovery experiment, as suggested by Reviewer #2, we ultimately concluded that the unlogged MAM does not adequately describe our modern analogue single-grain dose distributions.

To address this, we adopted a statistic that is more robust and better aligned with the structure of our single-grain dose data. Specifically, we now use the mode (the global maximum of the dose distribution) to represent the dominant well-bleached population of the distribution. To quantify *uncertainty in the modal estimate, we implemented a nonparametric bootstrap resampling procedure (n = 1000). In each iteration, a resample of equal size to the original dataset was drawn with replacement, and the mode was recalculated. The one-sigma (68%) confidence interval was then derived from the resulting bootstrap distribution of mode estimates. We believe this approach provides a more statistically robust and transparent representation of the dominant well-bleached component in our highly heterogeneously bleached modern fluvial samples. The updated results are shown in the revised Fig. 6 of the main text, and the new estimates are indeed much closer to a burial dose of 0 Gy than before. We have adjusted the corresponding interpretation and discussion accordingly.*

Minor comments:

Line 35: please change “remanent” to “remnant”.

Reply: Changed to “remnant”.

Line 42: Please add “e.g.,” before the list to indicate that it is illustrative.

Reply: Done.

Line 61-63: This sentence is misleading. As you mention later in line 66, the low residual doses reported in the cited studies are mainly due to the very low natural doses of the samples, rather than the pIRIR signals themselves. Please revise the sentence accordingly.

Reply: We have revised the sentence accordingly. Now it reads “Previous studies based on multi-grain measurements have demonstrated that low-temperature post-IR IRSL signals can be effectively bleached, resulting in residual doses, i.e., the luminescence dose remaining after laboratory bleaching, typically below 2 Gy (Reimann et al., 2011; Reimann and Tsukamoto, 2012).”

Line 65: The reference “Alexanderson and Murray, 2012” is cited in the text but is missing from the reference list.

Reply: “Alexanderson and Murray, 2012” is added to the reference list.

Line 83: Please clarify that the recuperation dose is related to the zero-dose measurement.

Reply: The text has been revised to indicate that the recuperation dose is derived from the zero-dose measurement.

Line 95: I suggest replacing “saturated” with “incompletely bleached” or writing it as “saturated, i.e., incompletely bleached”, since “saturated” could be interpreted as referring to old or naturally dose-saturated grains, whereas the intended meaning here is grains that have not fully reset their luminescence signal.

Reply: We had already mentioned “partially bleached” along with “saturated”. Thus, the intended meaning remains very clear in our opinion.

Line 113: Should this be “28° S”?

Reply: We understand the potential for confusion. The designation "27° S to 38° S" was intended to indicate the latitudinal extent of the Southern Central Andes, following conventions used in the literature, though we acknowledge that the southern boundary has been defined somewhat inconsistently across studies. We have now revised this to "27° S to 40° S" with appropriate reference support. In contrast, "28° S to 38° S" specifically refers to the latitudinal range of our studied catchments. To eliminate ambiguity, we have removed the term "Southern Central Andes" from the introduction and retained only the specific geographical descriptor "spanning 28° S to 38° S in Chile" to denote our study area.

Line 174: Please change this sentence to: "Growth curve fitting was performed using a general-order kinetic model (Guralnik et al., 2015), with the fit forced through the origin".

Reply: Done.

Line 175: What about recuperation as a criterion? Please provide a table summarising the number of grains measured, rejected, and accepted, together with the reasons for rejection (for both De and residual dose).

Reply: The application of recuperation as a rejection criterion requires careful consideration for modern samples with small natural signals. In such cases, recuperation values (R_0/N) are inherently high due to small denominators. Applying the recuperation-based rejection criteria to young samples would systematically exclude grains with smaller natural doses, precisely those representing better-bleached populations, resulting in biased distributions that overestimate burial doses. Thus, given that our samples are from modern fluvial analogues, we did not use recuperation as a rejection criterion for our analyses.

Two tables (Table S5 for D_e and Table S6 for residual dose measurements) summarising the number of grains measured, rejected, and accepted, together with the reasons for rejection, are now added to the supplementary document.

Line 182: I understand that performing dose recovery and residual dose tests at the single-grain level is very labor- and time-intensive. However, I wonder why smaller aliquots, such as 1 mm, were not used for these experiments, as these could be regarded as quasi–single-grain measurements.

Reply: We do not expect 1 mm aliquots to produce results that differ meaningfully from our 2 mm aliquot measurements. Reimann et al. (2012) showed that, because of the strong averaging effects in feldspars, even small-aliquot post-IR IRSL measurements containing ~20–30 grains are not an adequate substitute for true single-grain measurements. As a result, the added value of reducing the aliquot size from 2 mm to 1 mm for our samples is expected to be minimal.

Reference:

Reimann, T., Thomsen, K. J., Jain, M., Murray, A. S., and Frechen, M.: Single-grain dating of young sediments using the pIRIR signal from feldspar, Quat. Geochronol., 11, 28–41, https://doi.org/10.1016/j.quageo.2012.04.016, 2012.

Table 1: For consistency throughout the table, please add "at" to the preheat and IRSL, so they read "Preheat at 225 °C" and "IRSL at 50 °C".

Reply: Changes have been made to both Table 1 and Table S3.

Line 204: please clarify how many single grain discs were used for bleaching experiment.

Reply: We have now mentioned the number of single-grain discs measured for the bleaching experiment. We have added "Single-grain bleaching experiments were conducted on twenty discs (including 100 grains per disc) distributed across four samples, with 5 discs from CHLEA-3, 4 discs from CHLEA-7, 6 discs from CHLEA-10, and 5 discs from CHLEA-11" to the main text.

Line 215: Please remove the commas from "1,000" and "2,880 minutes" to read "1000 and 2880 minutes".

Reply: Done.

Fig. 4b: The IQR of Lx/Tx is shown without normalisation to the initial signal. In this case, the observed decrease in IQR may reflect the overall decay of the luminescence signal with bleaching time, rather than a real change in grain-to-grain variability. Please consider presenting either a normalised IQR (calculated from Lx/Tx values normalised to the initial signal) or a relative IQR (e.g. IQR(Lx/Tx) divided by the median Lx/Tx) to allow meaningful comparison across bleaching times.

Reply: Thank you for pointing this out and suggesting some alternatives. Following the suggestion, we have revised the figure (Fig. 4e) to show the IQR based on normalised L_x/T_x values to enable more meaningful comparisons.

Line 306: Please change "(Silverman, B.W., 1998)" to "(Silverman, 1998)".

Reply: Done.

Line 368: Please change "30,000 minutes" to "30000 minutes".

Reply: Done.

Fig. 6: I am wondering why the authors used an exponential fit rather than a linear fit. Please justify the choice of exponential fitting and discuss its impact on the extrapolated value.

Reply: We agree that our original fitting approach warranted reconsideration. The initial exponential fit was chosen because it maximized the R^2 value and captured the strongest apparent correlation between MAM D_e and mean residual dose.

However, since we now use the mode for both D_e and residual dose estimates, a linear model provides an equally good and appropriate description of the updated data. We have therefore refitted the relationship using a linear trendline and revised the corresponding discussion in Sections 4.1.4 and 4.3.1 of the main text.

Line 442: The citation "Ollerhead & Huntley" should be written as "Ollerhead and Huntley".

Reply: We have changed "Ollerhead & Huntley" to "Ollerhead and Huntley".

Line 467: Please mention that this risk of significant age overestimation is particularly important for young samples.

Reply: The following sentence has been added to the main text to address this point: "Additionally, the risk of significant age overestimation is particularly important for young samples."

Line 533: I think it is overstated to say different geomorphological settings here. Actually, all samples are from same geomorphological setting, with variation mainly in environmental gradients such as climate, slope, and lithology.

Reply: We have now rephrased the sentence to "The samples, sourced from catchments characterised by varying lithology, climate, and slope, were exposed to varying durations of light".

Supplementary:

Line 88: In the main text, only the acceptance criteria for single-grain measurements are mentioned.

Reply: We recognize the issue here. We have now changed the sentence to "DRRs were determined after subtracting the average residual dose from the average measured dose of all accepted aliquots, following the acceptance criteria analogous to those applied for single-grain measurements as outlined in Section 2.3 of the main text".

Line 88: The authors state that all pIRIR signals show good dose recovery ratios; however, this is not the case for sample CHLEA-6.

Reply: We have now changed the sentence to "CHLEA-11 showed good dose recovery results for all post-IR IRSL signals, while CHLEA-6 exhibited satisfactory dose recovery only for the post-IR $IRSL_{175}$ and post-IR $IRSL_{200}$ signals".

Table S3: Why is the single grain De protocol slightly different from the single grain dose recovery protocol? The single grain dose recovery does not include the IR LEDs at 200 °C at the end of the measurement.

Reply: This missing information has now been included in the caption of Table S3.

Line 258: "Table S7" should be corrected to "Table S8".

Reply: The table numbering has been revised to properly account for all tables included in the supplementary materials.

Line 291: "Table S8" should be corrected to "Table S10".

Reply: We have corrected the table numbering to account for additional tables included in the supplementary materials.