1-1-2017

Variability in quartz OSL signals caused by measurement uncertainties: Problems and solutions

Bo Li
University of Wollongong, bli@uow.edu.au

Zenobia Jacobs
University of Wollongong, zenobia@uow.edu.au

Richard G. Roberts
University of Wollongong, rgrob@uow.edu.au

Rex F. Galbraith
University College London, rex@stats.ucl.ac.uk

Jun Peng
Hunan University , Chinese Academy Of Sciences

Variability in quartz OSL signals caused by measurement uncertainties: Problems and solutions

Abstract

We simulated the variability in measured quartz optically stimulated luminescence (OSL) signals and dose response curves (DRCs) caused by measurement uncertainties, including counting statistics and instrumental irreproducibility. We find that these measurement errors can give rise to large variations in the observed luminescence signal and contribute to among-aliquot or among-grain scatter in DRCs and equivalent dose (De) values. Different measurement systems (i.e., luminescence readers) may have different counting statistics properties and, hence, may exhibit differing extents of variation in the observed OSL signal, even for the same sample. Our simulation shows that the random measurement uncertainties may result in some grains or aliquots being ¿saturated¿ (that is, the measured natural signal is consistent with, or lies above, the saturation level of the measured DRC) and that the rejection of these ¿saturated¿ grains may result in a truncated De distribution, with De underestimation for samples with natural doses close to saturation (e.g., twice the characteristic saturation dose, D0). We propose a new method to deal with this underestimation problem, in which standardised growth curves (SGCs) are established and the weighted-mean natural signal (Ln/Tn) from all measured grains is projected on to the corresponding SGCs to determine De. Our simulation results show that this method can produce reliable De estimates up to 5D0, which is far beyond the conventional limit of ¿2D0 using the standard SAR procedure.

Disciplines

Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Li, B., Jacobs, Z., Roberts, R. G., Galbraith, R. & Peng, J. (2017). Variability in quartz OSL signals caused by measurement uncertainties: Problems and solutions. Quaternary Geochronology, 41 11-25.

Variability in quartz OSL signals caused by measurement uncertainties: problems and solutions

Bo Li[1,*], Zenobia Jacobs[1, 2], Richard G. Roberts[1, 2], Rex Galbraith[3], Jun Peng[4]

[1] Centre for Archaeological Science, School of Earth and Environmental Sciences, University of Wollongong, Wollongong, NSW 2522, Australia
[2] ARC Centre of Excellence for Australian Biodiversity and Heritage, University of Wollongong, Wollongong, NSW 2522, Australia
[3] Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, UK
[4] School of resources, environment and safety engineering, Hunan University of Science and Technology, Hunan, China
*Corresponding author: bli@uow.edu.au

Abstract

We simulated the variability in measured quartz optically stimulated luminescence (OSL) signals and dose response curves (DRCs) caused by measurement uncertainties, including counting statistics and instrumental irreproducibility. We find that these measurement errors can give rise to large variations in the observed luminescence signal and contribute to among-aliquot or among-grain scatter in DRCs and equivalent dose (D_e) values. Different measurement systems (i.e., luminescence readers) may have different counting statistics properties and, hence, may exhibit differing extents of variation in the observed OSL signal, even for the same sample. Our simulation shows that the random measurement uncertainties may result in some grains or aliquots being 'saturated' (that is, the measured natural signal is consistent with, or lies above, the saturation level of the measured DRC) and that the rejection of these 'saturated' grains may result in a truncated D_e distribution, with D_e underestimation for samples with natural doses close to saturation (e.g., twice the characteristic saturation dose, D_0). We propose a new method to deal with this underestimation problem, in which standardised growth curves (SGCs) are established and the weighted-mean natural signal (L_n/T_n) from all measured grains is projected on to the corresponding SGCs to determine D_e. Our simulation results

show that this method can produce reliable D_e estimates up to $5D_0$, which is far beyond the conventional limit of ~$2D_0$ using the standard SAR procedure.

Keywords: counting statistics; standardised growth curves; instrumental irreproducibility; D_e underestimation

1. Introduction

Understanding differences in single-grain dose response curves (DRCs) is important since some studies have shown that D_e estimation can be dependent on the observed variation in the shape of the DRC (or characteristic saturation dose, D_0) (e.g., Gliganic et al., 2012; Duller, 2012; Li et al., 2016; Thomsen et al., 2016; Guo et al., 2017). Characterising the intrinsic variability of experimentally observed optically stimulated luminescence (OSL) signals from individual grains of quartz is, therefore, imperative to assess the reliability of DRCs and the resulting equivalent dose (D_e) values and ages.

A number of previous studies have investigated potential sources of variability in single-grain OSL signals and how they may affect D_e values. Observations typically included relate to: (a) grain-to-grain differences in the inherent luminescence sensitivity (signal brightness) of individual grains (e.g., Roberts et al., 1999; Duller et al., 2000; Jacobs et al., 2003, 2006); (b) grain-to-grain differences in decay curve shapes due to variance in the composition of the OSL signals as observed using continuous-wave (CW) stimulation (e.g., Roberts et al., 1999; Adamiec et al., 2000; Duller et al., 2000; Jacobs et al., 2003, 2006) and linearly-modulated (LM) stimulation (e.g., Singarayer, 2005; Jacobs et al., 2006, 2008); (c) differences in thermal stability of grains identified through pulsed-anneal measurements (e.g., Fan et al., 2011; Jacobs et al., 2016); (d) changes in decay curve shape during successive single-aliquot-regenerative-dose (SAR) cycles (e.g., progressive build-up of background or differential sensitisation of the various OSL components of the signal) (e.g., Jacobs et al., 2006, 2013; Gliganic et al., 2012); (e) checks for the extent of recuperation or thermal transfer of

OSL signals; and (f) OSL signals arising from different mineral grains or from grains with mineral inclusions that are optically sensitive (e.g., Jacobs et al., 2003; Duller, 2003). A set of objective rejection criteria (Jacobs et al., 2003; 2006) has been proposed to deal with many of the problems discussed above. But even after application of these criteria, and those proposed subsequently, significant overdispersion in D_e values remains under controlled laboratory conditions (e.g., in dose recovery experiments). It is likely that further intrinsic sources of variability affecting the OSL signal are present in samples of natural quartz, and that these may lead to the construction of variable or inaccurate DRCs.

Alternatively, or in addition, there may also be issues related to the error estimation procedures used to calculate the measurement uncertainties associated with the natural dose (L_n), regenerative-dose (L_x) and corresponding test-dose (T_n and T_x) signals used to construct the sensitivity-corrected (L_x/T_x) DRCs. The two main sources of measurement uncertainty include: (a) counting errors, and (b) instrument irreproducibility errors. Both of these error terms are propagated through every measurement of L and T. Counting error calculations usually assume that both the photon and dark counts detected by photomultiplier tubes follow a Poisson distribution (e.g., Galbraith et al., 1999; Galbraith, 2002), where the variance of the count equals the mean count (i.e., the variance-to-mean ratio (VMR) = 1). However, this assumption is usually invalid. Several studies have previously observed additional variance in the number of counts, such that the VMR is >1 (e.g., Galbraith et al., 1999, 2005; Li, 2007; Adamiec et al., 2012; Tudyka et al., 2016). The numerical simulation results of Bluszcz et al. (2015) suggest that the error associated with D_e values can be severely underestimated if a Poisson distribution is assumed. Furthermore, Adamiec et al. (2012) observed that different measurement systems may exhibit different degrees of additional variance in the photon counts and in the dark counts, and recommended that the uncertainty associated with each should be estimated independently for different measurement systems. Since calculation of the instrument irreproducibility error for a specific instrument includes counting error as a component that should be taken into account when estimating the irreproducibility-only error (e.g., Thomsen et al., 2005; Jacobs et al., 2006), the observations of Adamiec et al. (2012) also have a direct influence on estimation of that error. The application of ratio tests built into the SAR measurement sequence and used as rejection criteria (such as the recycling ratio, the OSL IR depletion ratio and the recuperation ratio) also require the accurate estimation of measurement uncertainties.

In this study, we explain our methods for estimating both the counting and instrument irreproducibility errors, and apply a series of numerical simulations to systematically examine the effect of these errors on the observed variability in OSL signals, including signal intensities, DRC shapes (and D_0 values) and estimation of D_e values. We also investigate how these measurement uncertainties may cause difficulties with D_e estimation for samples with natural doses close to the saturation level of the DRC when using a conventional SAR or standardised growth curve (SGC) procedure (Roberts et al., 1999; Murray and Wintle, 2000; Li et al., 2015a). To potentially overcome this problem, we propose a new method, based on construction of a SGC (Roberts and Duller, 2004; Li et al., 2015a; 2015b) and test the validity of this method using experimental data for a sample collected from an archaeological site in North Africa (Douka et al., 2014; Li et al., 2016).

2. Sample descriptions

Existing experimental data for a sediment sample (HF11) collected from the Haua Fteah Cave in Libya, were chosen to validate the numerical simulation results presented in this study. Details about the sample, and the collection, preparation and data analysis procedures are provided in Douka et al. (2014). They measured 1000 aliquots, each composed of quartz grains of 90–125 µm diameter, using standard single-grain discs with each grain-hole containing ~8 grains. They reported a weighted-mean D_e value of 126 ± 2 Gy (n=405) and a corresponding age of 66 ± 6 ka. Douka et al. (2014) also measured 1000 individual grains of 125–212 µm diameter, obtaining a weighted-mean D_e value of 131 ± 5 Gy (n=81) and an age of 71 ± 7 ka; the single- and multi-grain results are consistent at 1σ. Douka et al. (2014) made some pertinent observations about the OSL behaviour of the grains, including the following: (a) among the wide range of OSL decay curve shapes, some had much slower rates of decay than others, but carry-over of OSL signal between successive measurement cycles was not problematic; (b) the majority of DRCs could be fitted with a single saturating exponential function; (c) grains show a large range of DRC shapes; and (d) some of the natural OSL signals are close to, or in, dose saturation: 6.4% of the multi-grain aliquots and 13.1% of the single grains have L_n/T_n ratios that lie at, or above, the saturation level of the corresponding DRC and can be classified as either saturated grains or as Class 3 ('oversaturated') grains (Yoshida et al., 2000).

Li et al. (2016) re-analysed the multi-grain data for HF11. Based on the observation that fewer than 5% of the measured single grains contributed >80% of the total OSL signal, they deduced that the measured OSL signal from the multi-grain aliquots arises from only one or two grains, thereby effectively making these measurements surrogate single-grain measurements. Their analyses also confirmed the observations of Douka et al. (2014) that aliquots from the same and different samples exhibit a wide range of DRC shapes and D_0 levels. Importantly, Li et al. (2016) determined that the multi-grain aliquot DRCs could be divided into three broad groups (termed 'early', 'medium' and 'later') that saturated at different dose levels. The 'early' group saturated at low doses (<100 Gy), the 'later' group at much higher doses (>270 Gy) and the 'medium' group at intermediate doses. They found that each group of DRCs could be well-defined by a SGC (e.g., Roberts and Duller, 2004; Li et al., 2015a). The three SGCs were identical up to a dose of 50 Gy, above which they started to deviate significantly.

Li et al. (2016) calculated ages for each group using both full SAR DRCs for each multi-grain aliquot ('early' = 57 ± 6 ka, 'medium' = 70 ± 7 ka and 'later' = 70 ± 7 ka) and the SGC for each group (45 ± 4, 74 ± 7 and 71 ± 7 ka, respectively). They found that the SAR and SGC ages obtained for the 'early' group were significantly underestimated compared to those for the 'medium' and 'later' groups; they were also much younger than the ages obtained from the multi-grain and single-grain D_e values reported in Douka et al. (2014) and the age of 73 ± 5 ka based on multiple-elevated-temperature post-infrared infrared stimulated luminescence (MET-pIRIR) measurements of potassium-rich feldspar grains (Jacobs et al., 2017). For HF11, 64% (221 of 344) of the aliquots in the 'early' group were considered fully 'saturated' (i.e., the natural signal was consistent with, or lay above, the saturation level of the corresponding DRC); accordingly, finite D_e values for age determination could not be obtained for these aliquots. The ages for the 'medium' and 'later' groups are considered reliable: they are consistent with each other and with the OSL age reported in Douka et al. (2014) and the MET-pIRIR age (Jacobs et al., 2017). Only 3.5% of the aliquots in the 'medium' group were identified as fully saturated, and the 'later' group contained none.

3. Counting statistics

Adamiec et al. (2012) suggested that the uncertainty arising from counting statistics should be measured for individual measurement systems, because different instruments may have photon and dark counts that exhibit different amounts of variance. Building on these observations, Bluszcz et al. (2015) showed that, for their measurement systems, the photon and dark counts were best described by Negative Binomial (NB) distributions, instead of Poisson distributions. They proposed a method to correct for this variance in a luminescence signal calculated on the basis of a Poisson distribution, using a correction factor (K^2_{DEF}) determined as follows:

$$\mathrm{K}^2_{DEF} = (\mathrm{K}^2_{DC} - \mathrm{K}^2_{ph})\frac{B}{I} + \mathrm{K}^2_{ph} \tag{1}$$

where I is the signal (including both the photon counts and dark counts) detected by the photomultiplier, B is the dark count obtained by measuring a blank disc at room temperature and without any stimulation source, and K^2_{DC} and K^2_{ph} are the ratios between variance and mean values for the dark counts and photon counts, respectively (Adamiec et al., 2012). If the count data follow a Poisson distribution, the values of K^2_{DC} and K^2_{ph} are equal to unity, but if the count data exhibit additional variance then these values will be >1.

We used the method described in Adamiec et al. (2012) to determine the values of K^2_{DC} and K^2_{ph} for the luminescence system (Risø2) used to measure the multi-grain OSL signals for HF11. To estimate the dark count, a blank disc was held at room temperature (~20°C) and the counts recorded for 500 s without any light stimulation. For the photon counts, a blue filter pack, comprised of Schott BG39 and Corning 7-59 filters was placed in front of the photomultiplier and a constant photon flux was achieved by switching on the calibration light-emitting diode (LED) and measuring the counts for 500 s at room temperature.

Histograms and probability distributions of the dark and photon count rates are shown in Fig. 1a and 1b, respectively. The probability distributions are fitted using a negative binomial (NB) distribution function of the following form:

$$P(X = x) = \frac{\Gamma(k+x)}{x!\Gamma(k)}\left(\frac{k}{\mu+k}\right)^k\left(\frac{\mu}{\mu+k}\right)^x \qquad x = 0, 1, 2, 3, \ldots \qquad (2)$$

Where Γ represents the gamma function, x is the count number, k is a constant (the number of successful Bernoulli or binomial trials), and μ is the mean of the distribution. The variance of the NB distribution is $\mu + \mu^2/k$. The Risø2 dark counts are well described by a NB distribution (Fig. 1a), whereas the photon counts from the calibration LED are slightly negatively skewed. The estimated K^2_{DC} and K^2_{ph} values of 3.69 and 1.88, respectively, suggest that Risø2 has count data with greater variance than expected for a Poisson distribution. Correction factors should, therefore, be incorporated into the error calculation, based on eqn. (1), for all OSL signals measured using this system. We note that the K^2_{DC} and K^2_{ph} values for Risø2 are similar to those obtained for 'Eiger' at the University of Bern (Adamiec et al., 2012), but they are higher than the values obtained for the other two readers at that laboratory; we observed a similar range of values for the four measurement systems tested in our laboratory.

4. Instrumental irreproducibility

Instrumental irreproducibility is an estimate of all variability in OSL signals arising solely from the instrument; this includes variability associated with heating, light stimulation, movement of discs between successive measurements, and repositioning of the laser for single-grain measurements. The uncertainty associated with instrument irreproducibility is assumed to be the same for different samples measured on the same instrument. The instrument irreproducibility error associated with the measurement of single grains of quartz using the green laser attachment on Risø systems has been investigated previously (Duller et al., 1999; Truscott et al., 2000; Thomsen et al., 2005; Jacobs et al., 2006). These studies used slightly different approaches, but in essence instrument irreproducibility was determined by repeatedly irradiating, preheating and optically stimulating the same grain (e.g., 10 times or more) to obtain a set of L_x values or L_x/T_x ratios. The variance of the latter (S^2) was expected to be the sum of the variances for instrumental irreproducibility and counting statistics, so the former (${\sigma_{ins}}^2$) could be estimated using the following equation:

$$ {\sigma_{ins}}^2 = S^2 - {\sigma_{CS}}^2 \quad (3) $$

where ${\sigma_{CS}}^2$ represents the variance arising from counting statistics. Relative standard errors for instrument irreproducibility of between about 2.5 and 3.5% per OSL measurement have been reported for single-grain quartz OSL measurements (e.g., Truscott et al., 2000; Thomsen et al., 2005; Jacobs et al., 2006). As the calculation of this value, however, is dependent on the error arising from counting statistics (Eqn. 3), then the estimate of instrument irreproducibility may be incorrect if the dark and photon counts for the particular measurement system are assumed to have a Poisson distribution but are, in fact, more dispersed; the effect will be particularly acute when the luminescence sensitivity of the grains is relatively low and, the OSL counts are close to background.

5. Numerical simulation

5.1. Description of simulation method

The main aim of this stimulation is to test the effect of σ_{CS} and σ_{ins} on the scatter of experimentally observed OSL signals. We used a similar method to that proposed by Bluszcz et al. (2015) to generate pseudo-random counts, using the built-in random number generation function in R (R Core Team, 2016). Fig. 2 is a summary flowchart of the steps involved in the simulation, which involves the following steps:

1) Fit experimental single-grain T_n data with a gamma function. We first quantified the luminescence sensitivity (inherent grain brightness) distribution of sample HF11 to use data from a real sample as the basis for our simulation. We used T_n (the net OSL signal from a test dose of ~8.5 Gy) to represent sensitivity; T_n was calculated from the OSL counts in the initial 0.2 s of optical stimulation (2 s in duration), minus a 'late light' background represented by the final 0.2 s. We then assumed a gamma distribution to describe the sensitivity data (following Cunningham et al., 2015) of the following form:

$$f(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)} \tag{4}$$

where x is the count number, α is a shape parameter and β is a scale parameter. Fig. 3 shows the probability distribution of T_n for a total of 734 aliquots of sample HF11. A wide range of sensitivities is observed, ranging from 134 to >30,000 cts/0.2 s. The distribution is well-described by a gamma distribution (red line in Fig. 3), with α and β values of 2.3217 and 2.5843, respectively.

2) Generate of single-grain OSL sensitivities from a gamma distribution. The luminescence sensitivity of the ith modelled grain (η_i) is generated by randomly drawing from the gamma distribution obtained in Step 1. This value is considered the 'true' sensitivity of the grain.

3) Generate OSL signals based on a pre-determined DRC function The standard SAR procedure is modelled by generating a series of OSL signals for a range of doses (including 'natural' [L_n(i)], 'regenerative' [L_x(i)] and 'test dose' [T_n(i) and T_x(i)] signals), based on an assumed DRC function. To simplify the model, we assumed that the DRC follows a single saturating exponential function of the

form $Y = A[1–\exp(–X/D_0)]$, where Y is the test-dose corrected signal, X is the regenerative dose, D_0 is the characteristic saturation dose, and A is a constant. We also assumed that there is no sensitivity change or thermal transfer/recuperation between successive OSL measurements. Each of the OSL signals (L_x or T_x) is represented by 3 components: dark counts (B), fast-decaying signal (I_f) and slow-decaying signal (I_s). B is constant throughout all OSL measurements, and is determined independently (see section 3 and Fig. 1). I_f is assumed to be fully bleached during each OSL measurement, so it can be modelled according to the pre-determined DRC function, which can be described as follows:

$$I_f(i) = \eta_i \frac{1-\exp\left(-\frac{D}{D_0}\right)}{1-\exp\left(-\frac{D_t}{D_0}\right)} \tag{5}$$

where η_i is the sensitivity of the ith grain, D is the 'natural' or 'regenerative' dose, D_t is the test dose, and D_0 is the characteristic saturation dose. I_s is dose-dependent and assumed to decay negligibly during each OSL measurement; accordingly, it cannot be modelled using a DRC function. To model the contribution of I_s, we investigated the experimental data of I_f and I_s for sample HF11. We found a positive correlation between I_f and I_s, with the majority of I_s/I_f ratios falling in the range 0–0.05 (Fig. 4). To estimate the value of I_s, we then multiplied I_f by the median value (0.024) of the I_s/I_f ratios. This method, however, predicts a negligible I_s when I_f is small (e.g., only a few hundred counts) and this is not true, especially for zero-dose signals that are dominated by the slow-decaying component. To avoid this problem, we added a constant count rate of 70 cts/0.2 s to all modelled values of I_s, based on the minimum experimental values of I_s for the HF11 aliquots.

4) Add σ_{CS} to the OSL signals. We assumed that both the dark counts and photon counts follow a NB distribution. However, the B, I_f and I_s distributions have to be generated from separate NB distributions with different values for the mean and variance. The dark count numbers can be drawn from a NB distribution with mean B and variance $K_{DC}^2 B$, while the count number for I_f has mean I_f and variance $K_{ph}^2 I_f$, and the count number for I_s has mean I_s and variance $K_{ph}^2 I_S$. The initial OSL counts (L_i) for the i^{th} grain can, thus, be obtained using the following formula:

$$L_i(i) = I_f(i) + I_s(i) + B_i \quad (6)$$

where $I_f(i)$, $I_s(i)$ and B_i are drawn from their corresponding distributions. The variance of L_i can then be estimated as:

$$\mathrm{Var}(L_i) = K_{ph}^2(I_f + I_s) + K_{DC}^2 B \quad (7)$$

Similarly, the background signal (L_B) for the i[th] grain can be obtained using the following formula:

$$L_B(i) = I_s(i) + B_i \quad (8)$$

and its variance can be estimated as:

$$\mathrm{Var}(L_B) = K_{ph}^2 I_s + K_{DC}^2 B \quad (9)$$

We emphasise that the initial signal (L_i) and background signal (L_B) must be estimated separately to allow for variations caused by counting error.

5) <u>Add σ_{ins} to the net OSL signal</u>. Once the initial and background signal counts have been generated in Step 4, the net OSL signal is generated by drawing from a normal distribution with mean equal to $L_i(i) - L_B(i)$ and a relative standard deviation equal to the assigned σ_{ins} (e.g., 0.02 or 2%). The standard error of the net OSL signal is then estimated as:

$$\sqrt{\mathrm{Var}(L_i) + \mathrm{Var}(L_B) + (L_i - L_B)^2 \sigma_{ins}^2} \quad (10)$$

6) <u>Construct DRCs using the sensitivity-corrected OSL signals</u>. The standard SAR procedure is then simulated to generate a series of L_n, L_x and T_n and T_x values using the method described in Steps 2–5. The L_x/T_x ratios, and their associated uncertainties are then calculated.

7) Steps 2–6 are repeated a number of times (e.g., n = 500) to simulate a sediment sample containing grains with different OSL sensitivities.

5.2. Simulation of DRCs

We simulated DRCs to quantify variability in D_0 values and DRC shape as a result of σ_{CS} and σ_{ins}. In this simulation, we used a representative dose sequence similar to that used for HF11 by Douka et al. (2014); the simulation sequence consisted of six regenerative doses at 1.5 (in place of a zero-dose cycle), 30, 67.5, 120, 180 and 270 Gy, a repeat dose at 120 Gy, and a fixed test dose of 8.5 Gy. We modelled the DRCs using the K_{ph}^2 and K_{DC}^2 values for three different measurement systems—Risø2 (presented in this study) and 'Ermintrude' and 'Moench' (reported by Adamiec et al., 2012) (Table 1). For each measurement system, we assumed a constant dark count rate of 15 cts/0.2 s (based on Risø2) and simulated the DRCs for four combinations of D_0 and σ_{ins}: (1) D_0 = 50 Gy and σ_{ins} = 2%; (2) D_0 = 50 Gy and σ_{ins} = 4%; (3) D_0 = 200 Gy and σ_{ins} = 2%; and (4) D_0 = 200 Gy and σ_{ins} = 4%. So, the regenerative dose range corresponds to 5.4D_0 (combinations 1 and 2) or 1.35D_0 (combinations 3 and 4).

Table 1 summarises the simulation data and results for all four D_0 and σ_{ins} combinations and three measurement systems. The left-hand panels in Fig. 5 shows the simulated L_x/T_x ratios for 500 grains at different regenerative doses for each of the four combinations of D_0 and σ_{ins} using the K_{ph}^2 and K_{DC}^2values for Risø2. The red line in each plot represents the common DRC when the data points for all 500 grains are fitted with a single saturating exponential function. The right-hand panels in Fig. 5 are histograms of D_0 values calculated for each individual DRC (n = 500 in each panel) for each of the four simulation combinations. The same data sets are presented for the other two measurement systems in Fig. S1 and S2.

We make the following two important observations: (a) there are significant grain-to-grain variations in the L_x/T_x ratios; and (b) there are also significant grain-to-grain variations in the D_0 values calculated from individual DRCs. The latter range between about 35 and 70 Gy for the 50 Gy D_0 simulations (Fig. 5b,d), with ranges of 140–330 Gy (Fig. 5f) and 130–380 Gy (Fig. 5h) for the 200

Gy D_0 simulations using σ_{ins} = 2% and 4%, respectively. Although the mean D_0 values are consistent with the applied D_0 values in all simulated scenarios, the standard deviations increase with an increase in σ_{ins} and also with an increase in D_0.

We also observe that the common DRCs (red lines in Fig. 5, S1 and S2) have mean D_0 values that are indistinguishable from 50 or 200 Gy, and that all three measurement systems have similar extents of grain-to-grain scatter in the L_x/T_x ratios and corresponding D_0 values. This indicates that the main source of variability in the DRCs for the simulated sample arise from σ_{ins}, which is probably because most of the simulated grains (based on the experimental data from sample HF11) have bright signals (Fig. 1), so σ_{CS} is relatively small. For samples that contain a larger proportion of dim grains, σ_{CS} may contribute significantly to scatter, because of the relatively larger contribution from dark counts to the observed OSL signal. For some readers (e.g., Risø2 and Ermintrude), the K_{DC}^2 values are comparatively larger, thus exerting a relatively larger influence on DRC shapes and the spread in D_0 values. The simulation results also show that DRC shapes and the corresponding D_0 values are significantly affected by the range of regenerative doses (i.e., the maximum regenerative dose) used for construct the DRCs. The wider range of simulated D_0 values obtained for D_0 = 200 Gy is likely due, at least in part, to the restricted range of regenerative doses compared to the true D_0 value; that is, the maximum regenerative dose applied (270 Gy) is only 1.35D_0. Measuring higher regenerative doses may allow the true D_0 value to be better constrained.

5.3. Simulation of D_e values

We have demonstrated that the variability in D_0 values in our simulations can be explained by differences in σ_{CS} and σ_{ins}. We now need to determine how this variability in D_0 value and DRC shape might affect the accuracy of D_e estimates. To do so, we used the same method described above to model a range of surrogate 'natural' doses (P) and corresponding DRCs based on a simulated SAR sequence. We modelled P values ranging from $0.3D_0$ up to $5D_0$; that is, if the D_0 value is 50 Gy, then we simulated 500 grains at P values of between 15 Gy ($0.3D_0$) and 250 Gy ($5D_0$). For each chosen P value, the sensitivity distribution of 500 grains was randomly generated from the gamma distribution shown in Fig. 3, with the same distribution used for each group of grains. To mimic the standard SAR procedure, each grain was also given 7 regenerative doses scaled to the size of P (i.e., 0.01P, 0.2P, 0.45P, 0.8P, 1.2P and 1.8P, with a repeat dose at 0.8P) and a test dose of 8.5 Gy. σ_{CS} and σ_{ins} were added to each of the signals (section 5.2). Individual DRCs were fitted and D_e values estimated for each grain using the built-in function `calSARED()` provided in the R-package 'numOSL' (Peng et al., 2013; Peng and Li, 2017).

Simulated D_e values for 500 grains at each of four P values (50, 100, 150 and 200 Gy) are shown as radial plots in Fig. 6a–d. These results are based on the K_{ph}^2 and K_{DC}^2 values for Risø2, σ_{ins} = 2% and a D_0 value of 50 Gy, so that P = 50 Gy represents $1D_0$, P = 100 Gy, 150 Gy and 200 Gy represent $2D_0$, $3D_0$ and $4D_0$, respectively. All grains with the lowest P (50 Gy) yielded finite D_e values, and most of these (~97%) are consistent with P at 2σ. The weighted-mean D_e value of 49.8 ± 0.2 Gy, calculated using the central age model (CAM; Galbraith et al., 1999), is indistinguishable from P (Fig. 6a). For a P value of 100 Gy, all but 3 of the grains yielded finite D_e values; the resulting CAM D_e value of 97.6 ± 0.5 Gy only underestimates P slightly (Fig. 6b). For the larger P values, 150 and 200 Gy, which correspond to $3D_0$ and $4D_0$, only 64% and 30% of the grains yielded finite D_e values, respectively (Table 1) and these gave CAM D_e values that are significantly smaller than P (by around 13% and 30%, respectively) (Fig. 6c,d).

A compounding effect of variability in DRCs due to σ_{CS} and σ_{ins} is the increased likelihood that, at doses much greater than $2D_0$, L_n/T_n may sometimes intercept the DRC and sometimes not. This will lead to D_e distributions that can be described as 'truncated', so that only the leading edge of a distribution of D_e values (i.e., the finite D_e values) is included in the weighted-mean D_e value for a sample. Such truncated distribution will give rise to an underestimation of P, even if all grains share the same DRC or D_0 values as in this simulation. Fig. 7 shows the L_n/T_n ratios and corresponding DRCs for 4 simulated grains from the group with a D_0 value of 50 Gy and where P = 200 Gy ($4D_0$): two grains (#1 and #8) gave finite D_e values, whereas the other pair (#2 and #134) are fully saturated. The number of 'saturated' grains in each group with different natural doses (i.e., P = $2D_0$, $3D_0$, $4D_0$ and $5D_0$ respectively) are summarised in Table 1.

Fig. 8 shows the CAM D_e values (black circles) calculated using different combinations of D_0 (50 and 200 Gy) and σ_{ins} (2% and 4%), but the same K_{ph}^2 and K_{DC}^2 values (Risø2). The CAM D_e values are consistent with P up to $2D_0$, regardless of the size of D_0 or σ_{ins} (Fig. 8a–d). This is consistent with the conservative upper limit for D_e estimation suggested by Wintle and Murray (2006). Above >$2D_0$, the CAM D_e values systematically underestimate P and the degree of underestimation increases until a constant (maximum) CAM D_e value is attained. Using σ_{ins} = 2% results in a maximum CAM D_e value of ~140 Gy for a D_0 of 50 Gy, and ~530 Gy at D_0 = 200 Gy; these D_e values are about 30% and 34% smaller than the corresponding P values, respectively (Fig. 8a,c). The same pattern is observed when σ_{ins} is increased to 4%, except that the maximum CAM D_e value is smaller and the degree of underestimation of P is greater. Maximum CAM D_e values of ~120 Gy and ~490 Gy are obtained for grains with D_0 values of 50 and 200 Gy, respectively (Fig. 8b,d), representing a ~40% underestimation of P. Fig. S3 and S4 show that similar patterns in estimated D_e values are observed using the K_{ph}^2 and K_{DC}^2 values for Ermintrude and Moench.

6. A new method for D_e estimation

The simulation results suggest that the uncertainties associated with σ_{CS} and σ_{ins} can give rise to considerable variation in the shapes of measured DRCs (and in their D_0 values), the L_n/T_n ratios and, consequently, the D_e values, even though all grains in the simulation have common DRCs (and D_0 values) and P values. This variability poses a particular problem when L_n/T_n ratios are >$2D_0$, as the L_n/T_n ratio for some grains may be consistent with, or lie above, the saturation level of the corresponding measured DRC (e.g., grains #2 and #134 in Fig. 7). These 'saturated' grains will yield infinite D_e values and, hence, be rejected from the final D_e estimation, resulting in truncation of the full D_e distribution and an underestimation of the sample D_e (assuming that all grains share the same DRC or D_0 value).

To circumvent the problem associated with saturation of some grains above $2D_0$ and the truncation of the D_e distribution, we propose a new method for D_e estimation based on the full ('untruncated') distribution of L_n/T_n ratios for all aliquots or grains. This method builds on previous methods to establish SGCs (Roberts and Duller, 2004; Li et al., 2015a, 2015b, 2016), and can be divided into several steps. The first three steps are similar to the SGC D_e estimation procedure of Li et al. (2016), but Steps 4 and 5 are new to the method proposed here:

1) Apply the SAR procedure to individual grains or aliquots to calculate L_n/T_n and L_x/T_x ratios.

2) Separate the grains or aliquots into three different groups ('early', 'medium' and 'later') based on their relative saturation characteristics, so that grains or aliquots in the same group share a common DRC. This can be achieved by using the L_x/T_x ratios calculated for two different regenerative doses (Li et al., 2016).

3) Establish SGCs for the three groups, using the least squares (LS)-normalisation procedure of Li et al. (2016), which involves the following steps: a) fit L_x/T_x ratios for all grains or aliquots using a best-fit model; b) re-normalise the L_x/T_x ratios for each grain or aliquot using scaling factors that minimise the difference between the re-scaled L_x/T_x ratios and the fitted DRC; as each grain is treated individually, different scaling factors are determined for each grain; and (c) repeat the fitting and re-normalisation procedures iteratively until there is negligible change in the relative standard deviation of the re-normalised L_x/T_x ratios. This LS-normalisation procedure can be implemented using the `lsNORM()` function in the R package 'numOSL' (Peng et al., 2013; Peng and Li, 2017).

4) Re-normalise the L_n/T_n ratios for individual grains or aliquots. For those measured using a full SAR cycle, re-normalisation can be achieved by multiplying L_n/T_n by the scaling factors

determined in Step 3, to establish the SGC. Grains and aliquots for which only L_n/T_n and one additional regenerative-dose signal (L_r/T_r) were measured can be re-normalised using the following equation:

$$\frac{L'_n}{T'_n} = \frac{L_n}{T_n} \times \frac{f(D_r)}{L_r / T_r} \qquad (11)$$

Where L'_n/T'_n denotes the re-scaled L_n/T_n ratio, $f(x)$ denotes the SGC established by LS-normalisation, and D_r and L_r/T_r denote the additional regenerative dose and its corresponding sensitivity-corrected OSL signal, respectively.

5) Project the weighted-mean re-scaled L_n/T_n ratios for individual groups on to their corresponding SGCs to estimate the D_e value for each group.

In this new method, no grains are rejected because they are 'saturated'; apparent 'saturation' can arise simply from random errors associated with counting statistics and instrumental irreproducibility. By including all grains, a full and untruncated distribution of the re-normalised L_n/T_n ratios is obtained. As all grains (or aliquots) from the same group share the same DRC, and as we assume that all grains (or aliquots) have the same natural dose, then the distribution of their L_n/T_n ratios should be randomly distributed around a value corresponding to the natural dose (P).

We first tested the new method using the simulation data set presented in Section 5.3. The L_n/T_n and L_x/T_x ratios for 500 grains at four values of P (50, 100, 150 and 200 Gy) are shown in the left-hand panels of Fig. 9, while the right-hand panels show histograms of the distribution of L_n/T_n ratios. These results are based on the K_{ph}^2 and K_{DC}^2 values for Risø2, a D_0 value of 50 Gy and σ_{ins} = 0.02. A range of L_n/T_n ratios is obtained, even though all grains in a panel have the same P, distributed normally around a central value. We applied the CAM to calculate the weighted-mean L_n/T_n ratio for each P, and these are shown as horizontal lines in the left-hand panels of Fig. 9. To calculate the D_e value for each set of grains, the CAM L_n/T_n ratio is projected on to the best-fit DRCs (red lines). These D_e values are shown in Fig. 8 (as red squares) for different combinations of D_0 (50 and 200 Gy) and σ_{ins} (0.02 and 0.04). All D_e values based on the CAM L_n/T_n ratios are consistent with P at 2σ, even for P values as high as $5D_0$. We note that the size of the D_e uncertainties increases considerably P values $>4D_0$ (Fig. 8).

We also used the same simulation data set to estimate SGC D_e values for 500 individual grains at P values of 50, 100, 150 and 200 Gy, by projecting individual L_n/T_n ratios on to the best-fit SGCs (red lines in the left-hand panels of Fig. 9). The results are shown as blue triangles in Fig. 8. The SGC method yielded a similar pattern of D_e values to that obtained using standard SAR (black circles): reliable D_e values (i.e., indistinguishable from P), are obtained up to $2D_0$, but underestimation of D_e occurs when P increases relative to D_0. A larger underestimation in D_e is obtained from the SGC compared to SAR, consistent with previous observations of experimental data (Li et al., 2016).

7. Comparison with experimental data for HF11

To further test the new method, we applied it to the experimental OSL data collected for sample HF11. The aliquots from this sample have previously been divided into 'early', 'medium' and 'later' groups, according to the saturation characteristics of their DRCs (Li et al., 2016). Weighted-mean SAR D_e estimates of 108.1 ± 7.2, 133.6 ± 3.1 and 134.3 ± 4.1 Gy were calculated for the 'early', 'medium' and 'later' groups, respectively. The 'early' group underestimated the D_e significantly compared to the other two groups because a large proportion (~60%) of the aliquots was 'saturated' and the D_e distribution truncated (Li et al., 2016). The 'medium' and 'later' groups contained few saturated aliquots and their SAR D_e values were considered reliable. The left-hand panels in Fig. 10 show the re-normalised L_n/T_n and LS-normalised L_x/T_x ratios (blue squares and black circles, respectively) for the aliquots that comprise each of the three groups. The between-aliquot variation in the L_x/T_x ratios is similar to that observed in the simulation (Fig. 9), which implies that aliquots in the same group share a common DRC, so a SGC can be constructed for each group. There is, however, larger scatter in the re-normalised L_n/T_n ratios for the experimental data (Fig. 10), than in the simulated data (Fig. 9). This result is not unexpected, because a number of additional extrinsic sources of variability can influence a natural sample and such factors are not included in our simulation.

The distributions of re-normalised L_n/T_n ratios for each group of aliquots are shown as histograms (Fig. 10, middle panels) and radial plots (Fig. 10, right-hand panels). We used the CAM to estimate the weighted-mean re-normalised L_n/T_n values (horizontal dotted lines in the left-hand panels). These were projected on to the corresponding SGCs to obtain D_e values of 127.3 ± 5.8, 143.7 ± 3.1 and 134.2 ± 4.3 Gy for the 'early', 'medium' and 'later' groups, respectively. The latter D_e value is consistent at 1σ with the SAR D_e value for the 'later' group (134.3 ± 4.1 Gy) and the SGC D_e value for the 'medium' group is similar, albeit slightly larger at 2σ. The SGC D_e value for the 'early' group is also similar to the SGC D_e values obtained for the 'medium' and 'later' groups, because the use of re-normalised L_n/T_n ratios circumvents the problem of underestimation due to rejection of 'saturated' aliquots; this is consistent with our observations based on the simulations (Fig. 8).

The overdispersion values for the re-normalised L_n/T_n ratios are 12%, 9% and 15% for the 'early', 'medium' and 'later' groups, respectively. These values are considerably smaller than those obtained for the SAR D_e and SGC D_e distributions, which range from 26 to 44% (Li et al., 2016), presumably because of the non-linear relationship between OSL signal and dose. That is, a small change in the natural signal will produce a large change in D_e in the non-linear range of the DRC.

8. Discussion

Differences in the shape of DRCs for different grains or aliquots are often used to explain variations in their intrinsic physical properties (e.g., D_0 value). Our simulations demonstrate that the experimentally observed OSL signals and corresponding DRCs can be influenced significantly by measurement uncertainties (specifically σ_{CS} and σ_{ins}) and the measurement strategy used to determine the D_e values (such as the number and range of regenerative doses applied). To correctly characterise the intrinsic luminescence behaviours of different grains, measurement uncertainties must be taken into account appropriately. This includes the separate estimation of σ_{CS} for each measurement system, to establish whether the count data follow a Poisson distribution or exhibit additional variance. If the latter, then appropriate correction factors based on values of k^2_{DC} and k^2_{ph} should be applied to estimate the counting error associated with each luminescence signal (Eqn. 1); this will also influence estimation of σ_{ins}. Accounting for σ_{CS} and σ_{ins} is important for understanding the variability in observed luminescence behaviour, and also for correctly estimating the uncertainties associated with the measured OSL signals and resulting D_e values. Explanations of D_e distribution patterns and choices of appropriate age models are critically dependent on the correct estimation of the measurement errors and other sources of variation (Galbraith and Roberts, 2012).

For sediment samples, the sensitivity-corrected OSL signals and DRCs are subject to several sources of variation. Sensitivity changes may occur between successive luminescence measurements, due to the repeated laboratory application of heat and irradiation as part of the SAR procedure. Sensitivity changes between SAR cycles may result in different count numbers (and thus different σ_{CS} values), and changes in DRC shape if T_x is not strictly proportional to the preceding L_x signal. Furthermore, variable degrees of sensitivity change between measurement of L_x and T_x may result in large between-grain (or between-aliquot) scatter in L_x/T_x ratios, even though the grains (or aliquots) may have the same DRC or D_0 value. Li et al. (2015a,b) suggested that such scatter could be reduced by normalising of the DRCs using a single L_x/T_x ratio, the so-called 're-normalisation' method; the LS-normalisation procedure represents an improvement to this approach (Li et al., 2016). Thermal transfer may also play an important role in causing variations in the shape of DRCs, especially in the low-dose region of a DRC where the size of the thermally transferred signal is largest compared to the size of the regenerative dose signals. Finally, the composition of the quartz OSL signal (e.g., relative proportions of fast, medium and slow components) may also contribute to the variability observed in DRCs, given that different components of quartz OSL have been shown to differ considerably in their DRC shapes (e.g., Jain et al., 2003; Singarayer and Bailey, 2003).

In this study, our simulations are based on the simplifying assumption that there is no sensitivity change or thermal transfer between or within SAR measurement cycles, and that all grains have the same D_0 value and signal composition. The only between-grain variable incorporated in our simulation is luminescence sensitivity, based on the experimental data for sample HF11 (Fig. 3). So, although large variations in DRC shape were observed in the simulations (Fig. 5), this variability should be considered the minimum expected for a natural sample. Li et al. (2016) found that, even after LS-normalisation, the samples from Haua Fteah Cave in Libya, still had a ~2.5% variation in the sensitivity-corrected signal between different aliquots from the same DRC group. This remaining variability could not be explained by the measurement uncertainties, which suggests that additional sources of variation contribute to the observed scatter. Fortunately, variations of this magnitude do not prevent application of the SGC method, from which reliable estimates of D_e were obtained (Li et al., 2016).

An important outcome of our simulation is the demonstration that the variance associated with σ_{CS} and σ_{ins} may give rise to some L_n/T_n ratios consistent with, or higher than, the saturation level of the corresponding DRCs. The resulting truncated D_e distribution may yield an underestimate of the true D_e value (as only the leading edge of the D_e distribution is included) and this can be difficult to diagnose based only on the distribution patterns of D_e values (Fig. 6). Several methods have been proposed to deal with such samples. One approach is to rank grains according to their D_0 values and then calculate D_e values for grains with D_0 values that satisfy a particular criterion. It has been suggested that reliable D_e estimates can be obtained from the 'plateau' region in a plot of D_e against D_0 (e.g., Thomsen et al., 2016; Guo et al., 2017), but this method will only work if D_0 values are determined reliably and if grains differ in their true D_0 values. As demonstrated in our simulation (Fig. 5), the measured D_0 values can be highly variable, even when all grains have the same true D_0 value. Furthermore, not all DRCs can be fitted using a single saturating exponential function, so D_0 may not be comparable across all grains or aliquots. An alternative approach is to group aliquots or grains according to the ratio of the L_x/T_x ratios calculated for two different regenerative doses (Li et al., 2016); D_e values are then determined for the groups with the higher saturation doses, fewest saturated grains and consistent D_e values. This approach avoids reliance on D_0 as the selection criterion, but both methods require a large number of grains with sufficiently high saturation doses.

In this study, we propose a new method that includes all grains in the weighted-mean re-normalised L_n/T_n ratio, which is then projected on to the associated SGC. Based on numerical simulations, we show that this method can produce reliable D_e results well beyond the conservative limit of $2D_0$ (i.e., up to $4D_0$ or $5D_0$). We confirmed this finding by analysing experimental data from sample HF11. The D_0 value for the 'early' group of aliquots (~36 Gy; Li et al., 2016), prevented reliable D_e estimation beyond ~70 Gy (i.e., $2D_0$) using conventional SAR or SGC procedures. Using our new method, a D_e value of 127.3 ± 5.8 Gy (corresponding to ~$3.5D_0$) is obtained, demonstrating the potential of this method for dating samples with natural doses larger than $2D_0$. It is worth noting, however, that a large number of grains are required to produce a precise estimate of the weighted-mean L_n/T_n ratio and, thus, minimise the error in the calculated D_e value for samples approaching the saturation level of the SGC.

Age models, such as the CAM (Galbraith et al., 1999) have been used mostly in OSL dating for D_e estimation. Our simulation results show that the CAM appears to also work well with L_n/T_n ratios, although a firm statistical foundation for applying these age models to luminescence signals has yet to be established. Given the fact that the CAM is able to produce reliable estimates of the mean L_n/T_n ratio for well-bleached samples such as HF11, we anticipate that other age models (e.g., the minimum age model and finite mixture model; Galbraith and Roberts, 2012) may also be applicable to L_n/T_n ratios; to do so requires an appropriate overdispersion value for a well-bleached sample with the same mineral composition as the dated sample and, ideally, a similar age (Galbraith et al., 2005; Galbraith and Roberts, 2012).

Finally, the method of D_e estimation proposed here is based on the establishment of SGCs, so the reliability of D_e estimates based on weighted-mean L_n/T_n ratios relies heavily on constructing reliable

SGCs. For quartz OSL, there may be several groups of grains with different DRCs, so a SGC should be established for each group (Li et al., 2016). The combination of SGCs and weighted-mean L_n/T_n ratios not only allows D_e estimation beyond the conventional $2D_0$ limit for the standard SAR procedure, but can also save on instrument time; this is especially useful when dating a large number of samples with similar luminescence behaviours are dated, or when measuring a large number of grains or aliquots for each sample. But as SGC methods inevitably sacrifice useful information, such as the extent or efficacy of sensitivity correction (e.g., the recycling ratio), recuperation and OSL IR depletion, they should be used only after sample behaviour has been fully verified through comparisons with results obtained from full SAR measurements on a subset of grains or aliquots.

9. Conclusions

Counting statistics and instrumental uncertainties play important roles in the observed variability of measured luminescence signals and the shape of corresponding DRCs. Such variability depends in part on the measurement system used, because individual instruments can have different variances in relation to both counting statistics and instrumental irreproducibility. These measurement uncertainties may cause significant underestimates of D_e for samples with natural doses of $>2D_0$, due to the rejection of 'saturated' grains. The latter problem can be avoided by constructing SGCs and projecting the weighted-mean L_n/T_n ratio for all grains on to the corresponding SGCs. This enables reliable estimates of D_e to be obtained at doses well above the conventional limit of $2D_0$—conservatively up to $4D_0$ and possibly as high as $5D_0$. But further tests on known-age and well-bleached natural samples are needed to confirm the broader applicability of the approach provided here.

Acknowledgements

This research was funded by the Australian Research Council through Future Fellowships to Li (FT140100384) and Jacobs (FT150100138), and an Australian Laureate Fellowship to Roberts

(FL130100116). We thank Grzegorz Adamiec for his constructive comments on the manuscript. Yasaman Jafari and Terry Lachlan provided valuable support in the laboratory.

References

Adamiec, G., Heer, A.J., Bluszcz, A., 2012. Statistics of count numbers from a photomultiplier tube and its implications for error estimation. Radiation Measurements 47, 746–751.

Bluszcz, A., Adamiec, G., Heer, A.J., 2015. Estimation of equivalent dose and its uncertainty in the OSL SAR protocol when count numbers do not follow a Poisson distribution. Radiation Measurements 81, 46–54.

Cunningham, A.C., Wallinga, J., Hobo, N., Versendaal, A.J., Makaske, B., Middelkoop, H., 2015. Re-evaluating luminescence burial doses and bleaching of fluvial deposits using Bayesian computational statistics. Earth Surface Dynamics 3, 55–65.

Douka, K., Jacobs, Z., Lane, C., Grün, R., Farr, L., Hunt, C., Inglis, R.H., Reynolds, T., Albert, P., Aubert, M., Cullen, V., Hill, E., Kinsley, L., Roberts, R.G., Tomlinson, E.L., Wulf, S., Barker, G., 2014. The chronostratigraphy of the Haua Fteah cave (Cyrenaica, northeast Libya). Journal of Human Evolution 66, 39–63.

Duller, G.A.T., Bøtter-Jensen, L., Murray, A.S., 2000. Optical dating of single sand-sized grains of quartz: sources of variability. Radiation Measurements 32, 453–457.

Duller, G.A.T., 2003. Distinguishing quartz and feldspar in single grain luminescence measurements. Radiation Measurements 37, 161–165.

Duller G.A.T., 2012. Improving the accuracy and precision of equivalent doses determined using the optically stimulated luminescence signal from single grains of quartz. Radiation Measurements 47, 770–777.

Galbraith, R.F., Green, P.F., 1990. Estimating the component ages in a finite mixture. Nuclear Tracks and Radiation Measurements 17, 197–206.

Galbraith, R.F., Roberts, R.G., 2012. Statistical aspects of equivalent dose and error calculation and display in OSL dating: An overview and some recommendations. Quaternary Geochronology 11, 1–27.

Galbraith, R.F., Roberts, R.G., Laslett, G.M., Yoshida, H., Olley, J.M., 1999. Optical dating of single and multiple grains of quartz from Jinmium rock shelter, northern Australia, Part I: Experimental design and statistical models. Archaeometry 41, 339–364.

Galbraith, R.F., Roberts, R.G., Yoshida, H., 2005. Error variation in OSL palaeodose estimates from single aliquots of quartz: a factorial experiment. Radiation Measurements 39, 289–307.

Gliganic, L.A., Jacobs, Z., Roberts, R.G., 2012. Luminescence characteristics and dose distributions for quartz and feldspar grains from Mumba rockshelter, Tanzania. Archaeological and Anthropological Sciences 4, 115–135.

Guo, Y.J., Li, B., Zhang, J.F., Yuan, B.Y., Xie, F., Roberts, R.G., 2017. New ages for the Upper Palaeolithic site of Xibaimaying in the Nihewan Basin, northern China: implications for small-tool and microlithic industries in northeast Asia during Marine Isotope Stages 2 and 3. Journal of Quaternary Science, in press.

Fan, A.C., Li, S.H., Li, B., 2011. Observation of unstable fast component in OSL of quartz. Radiat Meas 46, 21-28.

Jacobs, Z., Duller, G.A.T., Wintle, A.G., 2003. Optical dating of dune sand from Blombos Cave, South Africa: II – single grain data. Journal of Human Evolution 44, 613–625.

Jacobs, Z., Duller, G.A.T., Wintle, A.G., 2006. Interpretation of single grain D_e distributions and calculation of D_e. Radiation Measurements 41, 264–277.

Jacobs, Z., Wintle, A.G., Roberts, R.G., Duller, G.A.T., 2008. Equivalent dose distributions from single grains of quartz at Sibudu, South Africa: context, causes and consequences for optical dating of archaeological deposits. Journal of Archaeological Science 35, 1808–1820.
Jacobs, Z., Hayes, E.H., Roberts, R.G., Galbraith, R.F., Henshilwood, C.S., 2013. An improved OSL chronology for the Still Bay layers at Blombos Cave, South Africa: Further tests of single-grain dating procedures and a re-evaluation of the timing of the Still Bay industry across southern Africa. Journal of Archaeological Science 40, 579–594.
Jacobs, Z., Jankowski, N.R., Dibble, H.L., Goldberg, P., McPherron, S.J.P., Sandgathe, D., Soressi, M., 2016. The age of three Middle Palaeolithic sites: Single-grain optically stimulated luminescence chronologies for Pech de l'Azé I, II and IV in France. Journal of Human Evolution 95, 80–103.
Li, B., 2007. A note on estimating the error when subtracting background counts from weak OSL. Ancient TL 25, 9–14.
Li, B., Roberts, R.G., Jacobs, Z., Li, S.H., 2015a. Potential of establishing a 'global standardised growth curve' (gSGC) for optical dating of quartz from sediments. Quaternary Geochronology 27, 94–104.
Li, B., Roberts, R.G., Jacobs, Z., Li, S.H., Guo, Y.J., 2015b. Construction of a 'global standardised growth curve' (gSGC) for infrared stimulated luminescence dating of K-feldspar. Quaternary Geochronology 27, 119–130.
Li, B., Jacobs, Z., Roberts, R.G., 2016. Investigation of the applicability of standardised growth curves for OSL dating of quartz from Haua Fteah cave, Libya. Quaternary Geochronology 35, 1–15.
Murray, A.S., Wintle, A.G., 2000. Luminescence dating of quartz using an improved single-aliquot regenerative-dose protocol. Radiation Measurements 32, 57–73.
Peng, J., Dong, Z.B., Han, F.Q., Long, H., Liu, X.J., 2013. R package numOSL: numeric routines for optically stimulated luminescence dating. Ancient TL 31, 41–48
Peng, J., Li, B., 2017. General R programs for equivalent dose determination based on single aliquot regenerative-dose (SAR) and standardised growth curve (SGC) methods. Quaternary Geochronology, submitted.
R Core Team, 2016. R: A language and environment for statistical computing. Vienna, Austria. URL: https://www.r-project.org/.
Roberts, H.M., Duller, G.A.T., 2004. Standardised growth curves for optical dating of sediment using multiple-grain aliquots. Radiation Measurements 38, 241–252.
Roberts, R.G., Galbraith, R.F., Olley, J.M., Yoshida, H., Laslett, G.M., 1999. Optical dating of single and multiple grains of quartz from Jinmium rock shelter, northern Australia, Part II: Results and implications. Archaeometry 41, 365–395.
Roberts, R.G., Galbraith, R.F., Yoshida, H., Laslett, G.M., Olley, J.M., 2000. Distinguishing dose populations in sediment mixtures: a test of single-grain optical dating procedures using mixtures of laboratory-dosed quartz. Radiation Measurements 32, 459–465.
Roberts, R.G., Jacobs, Z., Li, B., Jankowski, N.R., Cunningham, A.C., Rosenfeld, A.B., 2015. Optical dating in archaeology: Thirty years in retrospect and grand challenges for the future. Journal of Archaeological Science 56, 41–60.
Singarayer, J.S., Bailey, R.M., 2003. Further investigations of the quartz optically stimulated luminescence components using linear modulation. Radiation Measurements 37, 451–458.
Thomsen, K.J., Murray, A.S., Bøtter-Jensen, L., 2005. Sources of variability in OSL dose measurements using single grains of quartz. Radiation Measurements 39, 47–61.
Thomsen, K.J., Murray, A.S., Buylaert, J.P., Jain, M., Hansen, J.H., Aubry, T., 2016. Testing single-grain quartz OSL methods using sediment samples with independent age control from the Bordes-Fitte rockshelter (Roches d'Abilly site, Central France). Quaternary Geochronology 31, 77–96.

Truscott, A.J., Duller, G.A.T., Bøtter-Jensen, L., Murray, A.S., Wintle, A.G., 2000. Reproducibility of optically stimulated luminescence measurements from single grains of Al_2O_3:C and annealed quartz. Radiation Measurements 32, 447–451.

Tudyka, K., Adamiec, G., Bluszcz, A., 2016, Simulation of He+ induced afterpulses in PMTs. Review of Scientific Instruments 87, 063120.

Wintle, A.G., Murray, A.S., 2006. A review of quartz optically stimulated luminescence characteristics and their relevance in single-aliquot regeneration dating protocols. Radiation Measurements 41, 369–391.

Yoshida, H., Roberts, R.G., Olley, J.M., Laslett, G.M., Galbraith, R.F., 2000. Extending the age range of optical dating using single 'supergrains' of quartz. Radiation Measurements 32, 439–446.

Figure captions

Figure 1: Histograms of (a) dark counts in 1 s intervals (bin size = 10 s) and (b) calibration LED photon counts in 0.1 s intervals (bin size = 50 s). The grey bars/area show the negative binomial probability distributions with parameters fitted by the method of moments. Summary statistics are shown in each panel. Note that the negative binomial distribution fits the data well in (a) but not in (b), where the data are more negatively skewed.

Figure 2: Flowchart for simulation of OSL signals and corresponding DRCs.

Figure 3: Histogram of T_n signal intensity for 734 aliquots of sample HF11. The inset plot shows the cumulative density function (CDF) for these data, which are fitted by a gamma distribution (red lines) with the best-fit α and β values indicated.

Figure 4: Histogram showing the ratios between the slow-decaying signal (I_s) and fast-decaying signal (I_f) for 734 single aliquots of sample HF11. The inset plot shows the corresponding I_f and I_s signal intensities.

Figure 5: Simulated DRCs for a total of 500 grains in each panel, based on the K_{ph}^2 and K_{DC}^2 values for Risø2 and the four combinations of D_0 value and instrumental uncertainty (σ_{ins}): (a) D_0 = 50 Gy, σ_{ins} = 0.02; (c) D_0 = 50 Gy, σ_{ins} = 0.04; (e) D_0 = 200 Gy, σ_{ins} = 0.02; (g) D_0 = 200 Gy, σ_{ins} = 0.04. (b), (d), (f) and (h) show the distributions of D_0 values for individual simulated DRCs for the 500 grains in panels (a), (c), (e) and (g), respectively.

Figure 6: Radial plots showing the distributions of simulated D_e values for 500 grains at four surrogate natural doses (P values): (a) 50 Gy (b) 100 Gy (c) 150 Gy and (d) 200 Gy. The simulations used K_{ph}^2 and K_{DC}^2 values for Risø2, a D_0 value of 50 Gy and σ_{ins} value of 0.02. The black lines and grey shading in the radial plots represent the weighted mean of the data sets calculated using CAM and the associated ±2 standardised estimate band, respectively. The pink lines represent the P values.

Figure 7: Sensitivity-corrected natural signals (blue squares) and corresponding DRCs (red lines) for 4 simulated grains from the group with P = 200 Gy. These simulations are based on K_{ph}^2 and K_{DC}^2 values for Risø2, a D_0 value of 50 Gy and an instrumental uncertainty (σ_{ins}) of 2%.

Figure 8: Modelled SAR D_e values (black circles), SGC D_e values (blue triangles) and D_e values based on mean L_n/T_n ratios (red squares) plotted against the natural dose (P). The data in each panel are based on different combinations of D_0 (50 and 200 Gy) and σ_{ins} (0.02 and 0.04), but the same K_{ph}^2 and K_{DC}^2 values (for Risø2). The D_0 and σ_{ins} values used for these simulations are shown in each panel, with each data point based on the weighted-mean of 500 simulated grains; weighted-mean D_e values were calculated using the CAM.

Figure 9: Sensitivity-corrected natural (L_n/T_n, blue squares) and regenerative (L_x/T_x, black circles) ratios for 500 grains at four natural doses: (a) 50 Gy, (c) 100 Gy, (e) 150 Gy and (g) 200 Gy. The corresponding distributions of L_n/T_n ratios are shown in panels (b), (d), (f) and (h), respectively. Results are based on the K_{ph}^2 and K_{DC}^2values for Risø2, a D_0 value of 50 Gy and σ_{ins} value of 0.02.

Figure 10: Re-normalised L_n/T_n ratios and DRCs for different aliquots from the three groups (‘early’, ‘medium’ and ‘later’), recognised for sample HF11 (a,d and g). The distributions of re-normalised L_n/T_n ratios for the three groups are shown as histograms (b, e and h) and as radial plots (c, f and i). The grey shading in each of the radial plots represents the ±2 standardised estimate band, centred on the weighted-mean re-normalised L_n/T_n ratio, calculated using the CAM.

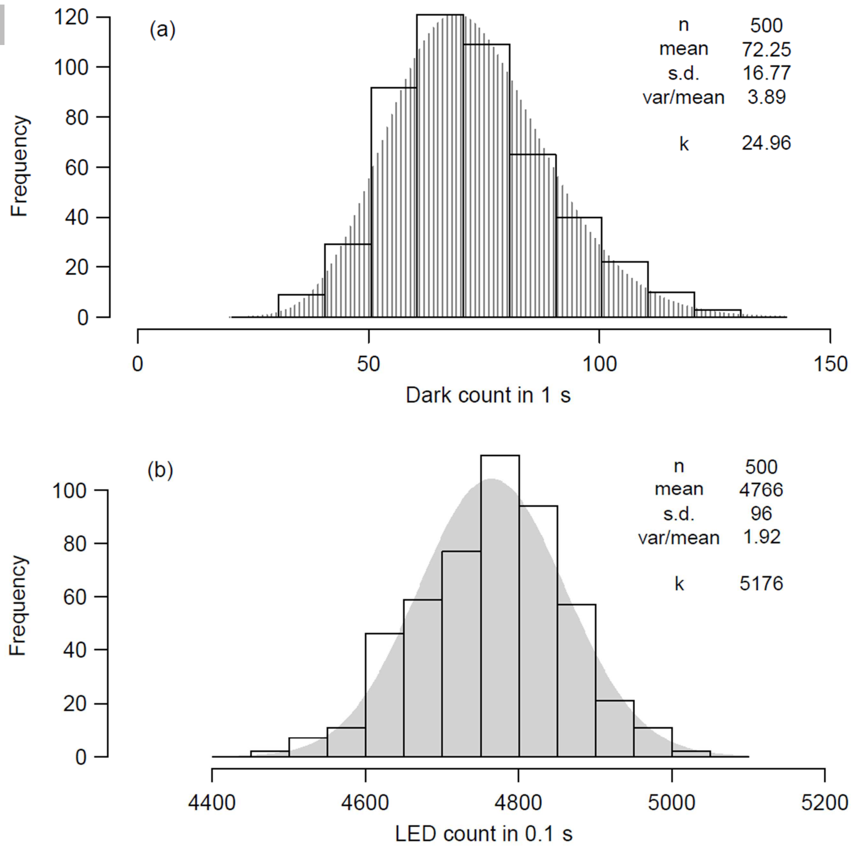

(a)
Frequency
n 500
mean 72.25
s.d. 16.77
var/mean 3.89
k 24.96
Dark count in 1 s
(b)
n 500
mean 4766
s.d. 96
var/mean 1.92
k 5176
LED count in 0.1 s

Figure 1

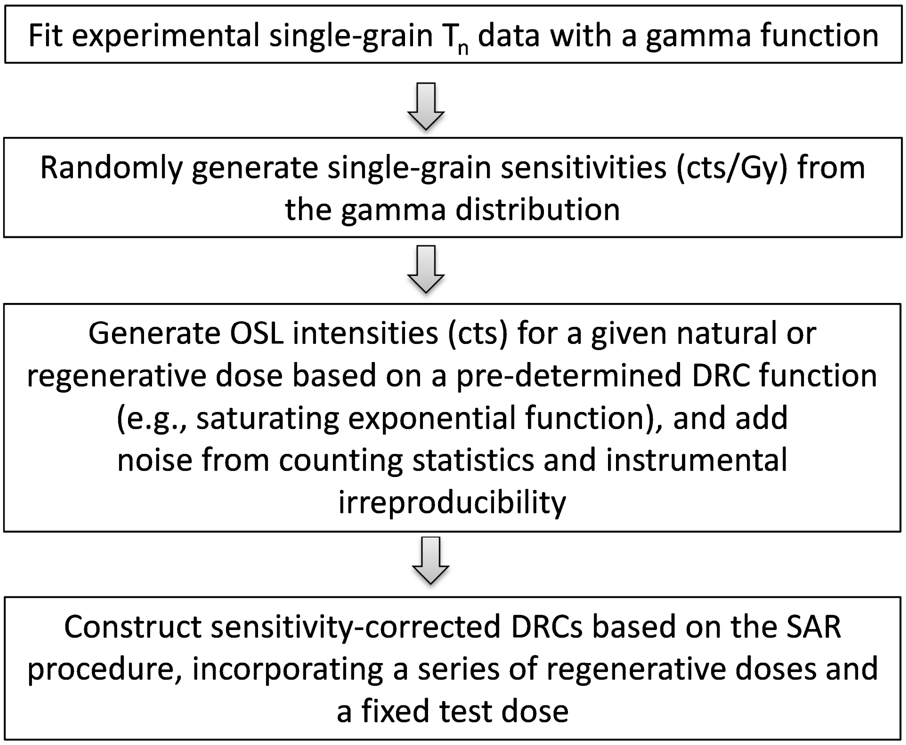

Fit experimental single-grain T_n data with a gamma function
Randomly generate single-grain sensitivities (cts/Gy) from the gamma distribution
Generate OSL intensities (cts) for a given natural or regenerative dose based on a pre-determined DRC function (e.g., saturating exponential function), and add noise from counting statistics and instrumental irreproducibility
Construct sensitivity-corrected DRCs based on the SAR procedure, incorporating a series of regenerative doses and a fixed test dose

Figure 2

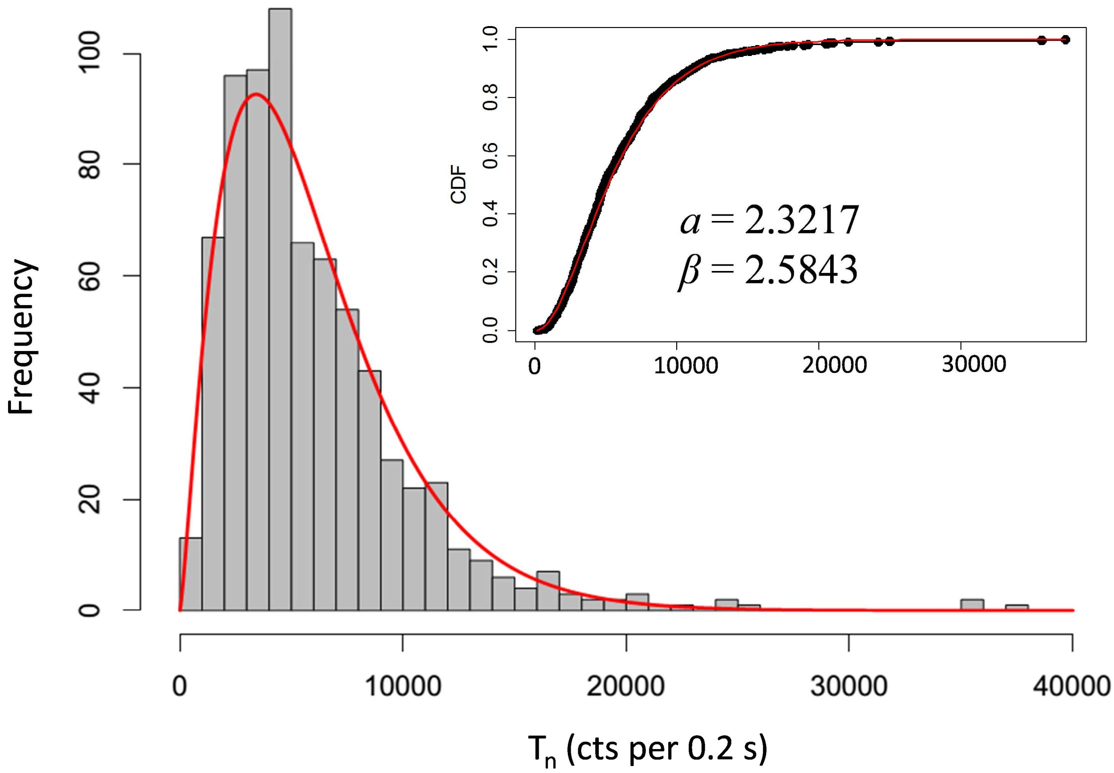
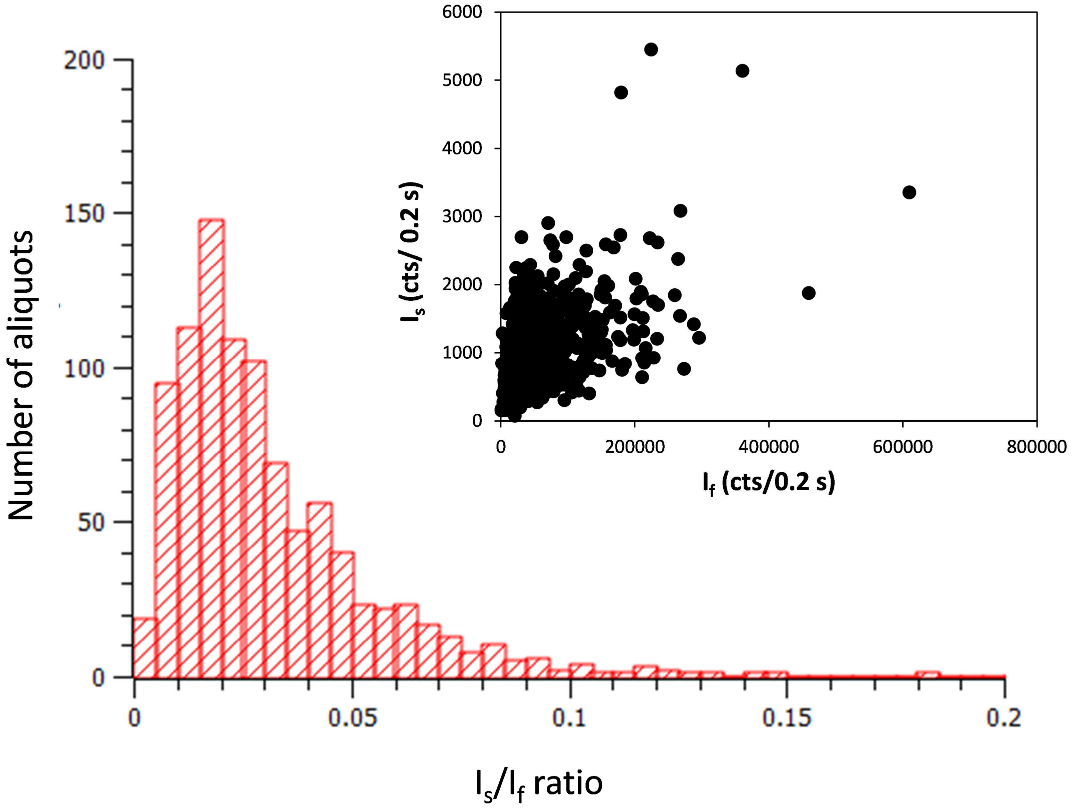

Frequency
0
20
40
60
80
100
0
10000
20000
30000
40000
T_n (cts per 0.2 s)
CDF
0.0
0.2
0.4
0.6
0.8
1.0
$a = 2.3217$
$\beta = 2.5843$

Figure 3



Number of aliquots
200
150
100
50
0
0
0.05
0.1
0.15
0.2
I_s/I_f ratio
6000
5000
4000
3000
2000
1000
0
I_s (cts/ 0.2 s)
0
200000
400000
600000
800000
I_f (cts/0.2 s)

Figure 4

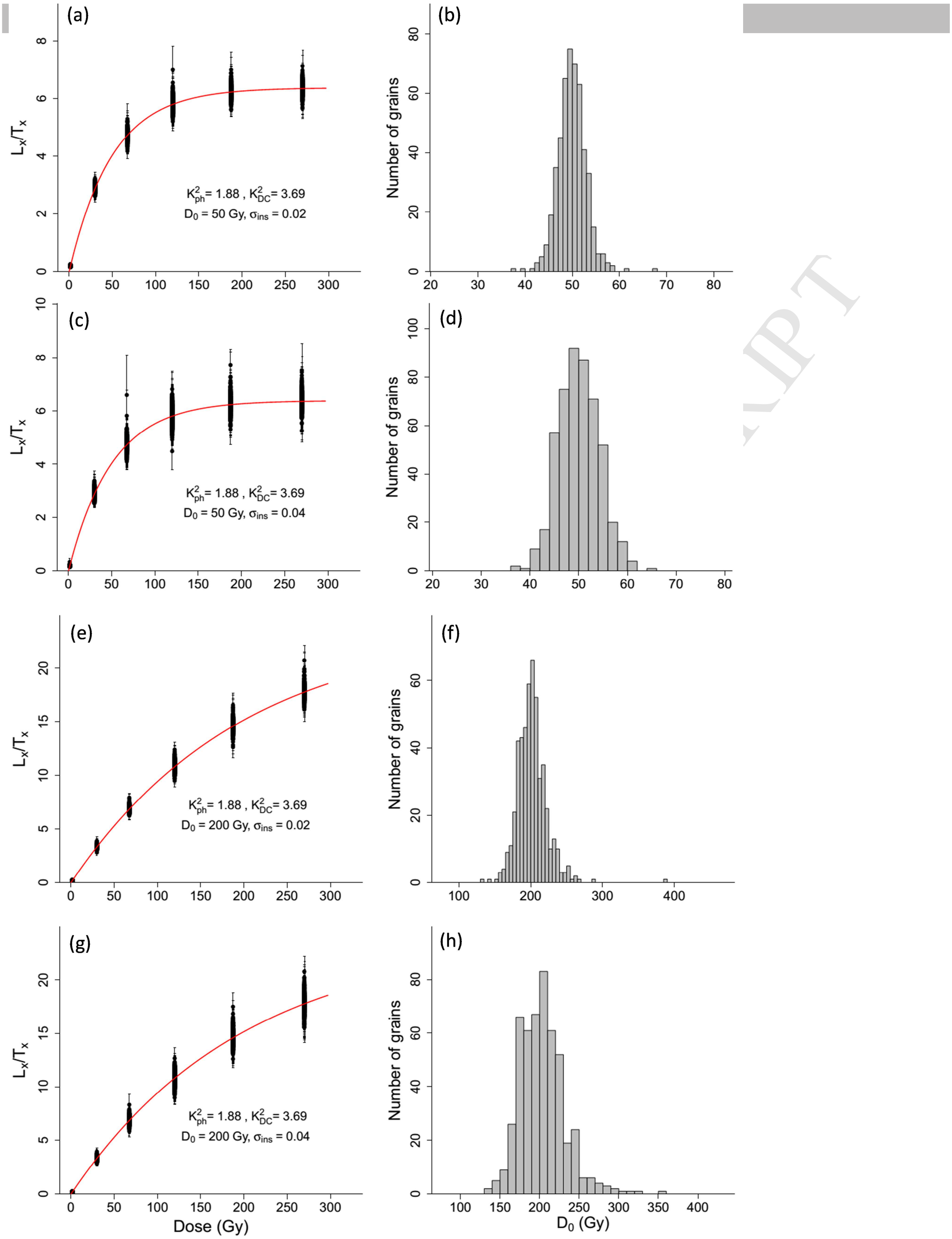
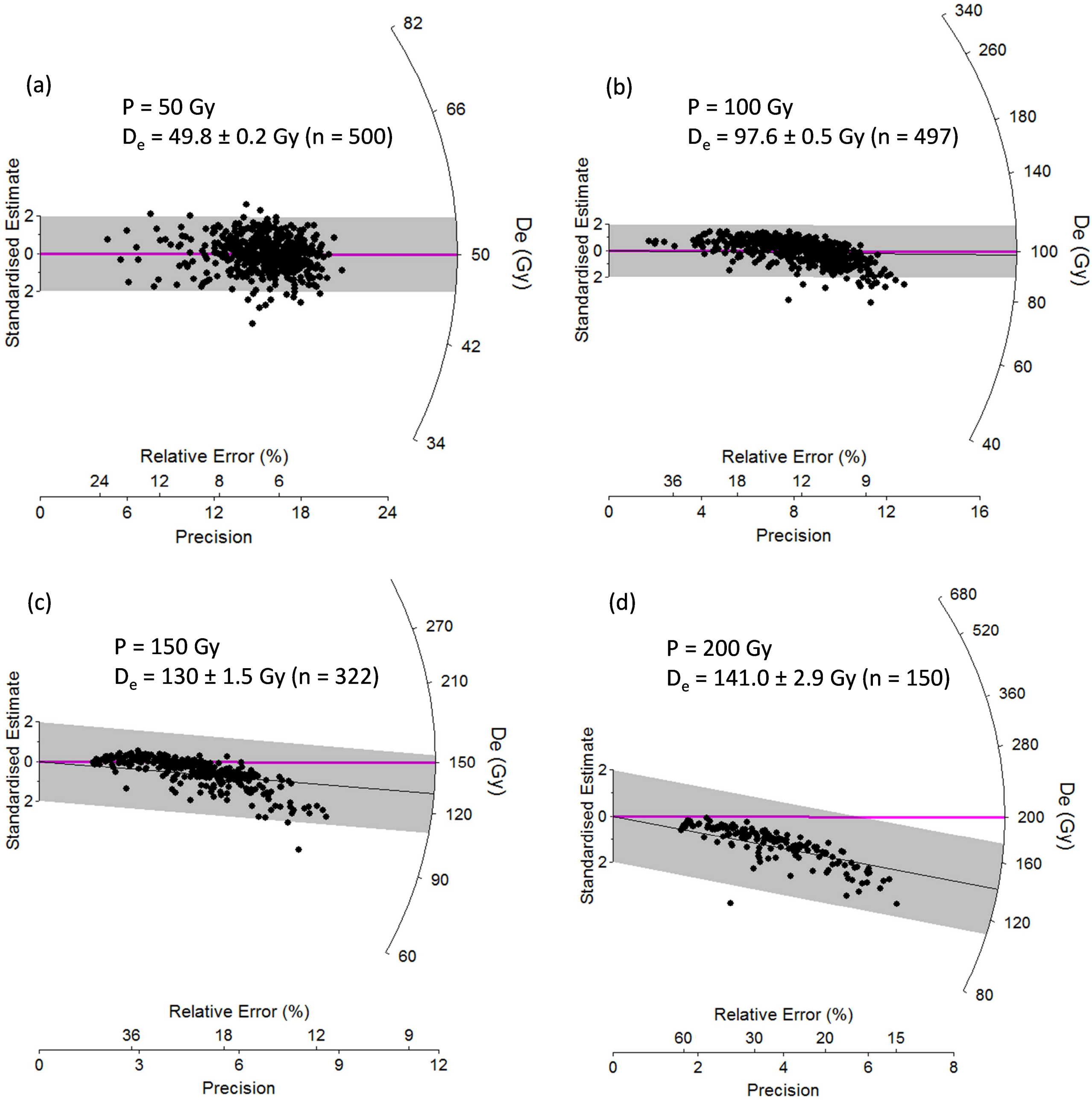

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
L_x/T_x
Number of grains
K^2_{ph}= 1.88 , K^2_{DC}= 3.69
D_0 = 50 Gy, σ_{ins} = 0.02
D_0 = 50 Gy, σ_{ins} = 0.04
D_0 = 200 Gy, σ_{ins} = 0.02
D_0 = 200 Gy, σ_{ins} = 0.04
Dose (Gy)
D_0 (Gy)

Figure 5



(a)
P = 50 Gy
D_e = 49.8 ± 0.2 Gy (n = 500)
(b)
P = 100 Gy
D_e = 97.6 ± 0.5 Gy (n = 497)
(c)
P = 150 Gy
D_e = 130 ± 1.5 Gy (n = 322)
(d)
P = 200 Gy
D_e = 141.0 ± 2.9 Gy (n = 150)
Standardised Estimate
Relative Error (%)
Precision
De (Gy)

Figure 6

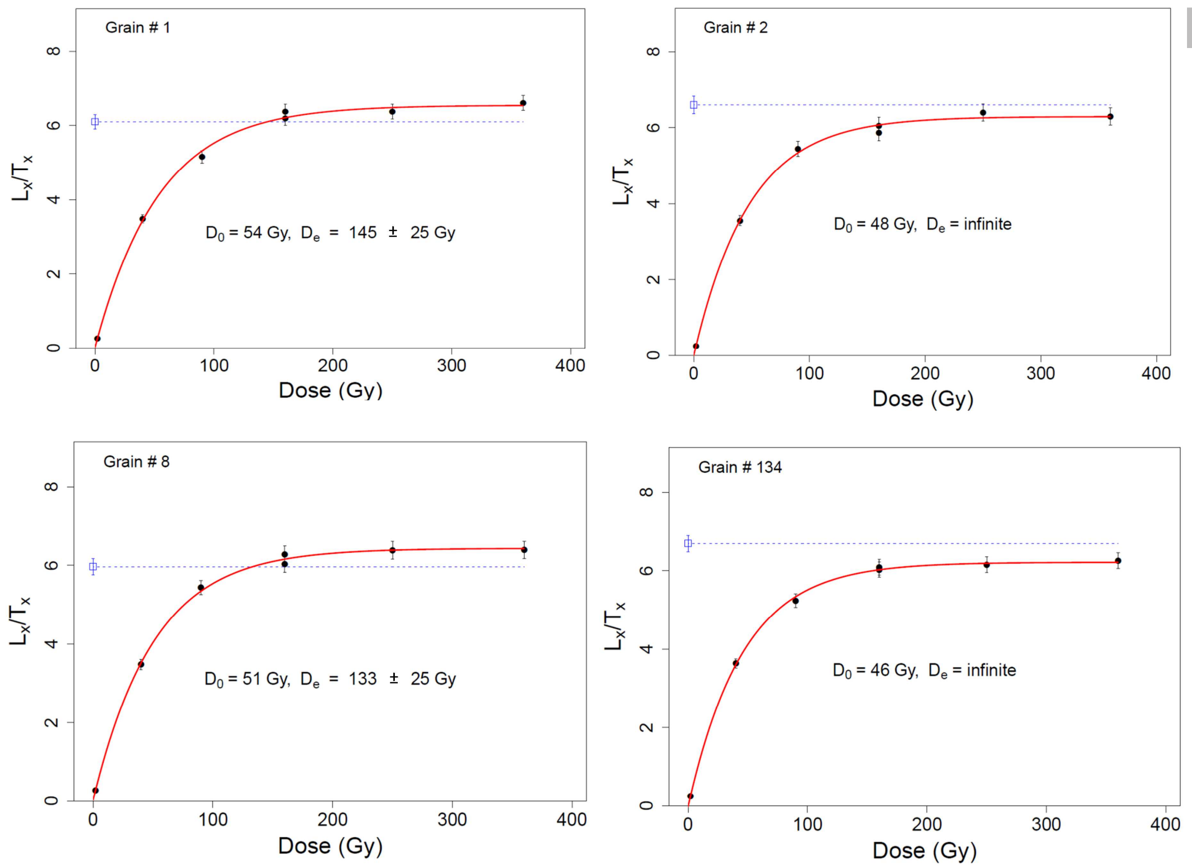

Grain # 1
D_0 = 54 Gy, D_e = 145 ± 25 Gy
Grain # 2
D_0 = 48 Gy, D_e = infinite
Grain # 8
D_0 = 51 Gy, D_e = 133 ± 25 Gy
Grain # 134
D_0 = 46 Gy, D_e = infinite
L_x/T_x
Dose (Gy)

Figure 7

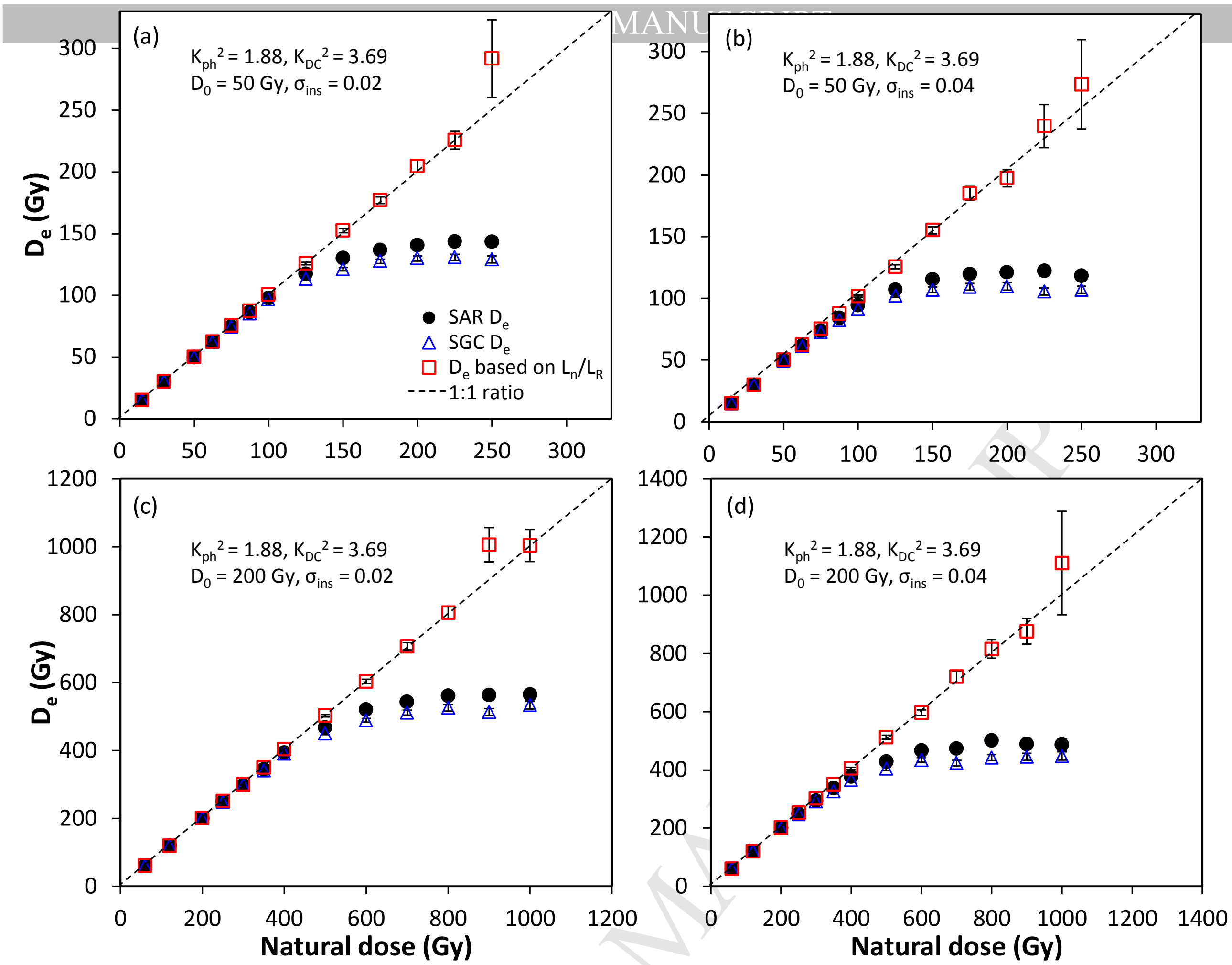

(a)
K_{ph}^2 = 1.88, K_{DC}^2 = 3.69
D_0 = 50 Gy, σ_{ins} = 0.02
SAR D_e
SGC D_e
D_e based on L_n/L_R
1:1 ratio
(b)
K_{ph}^2 = 1.88, K_{DC}^2 = 3.69
D_0 = 50 Gy, σ_{ins} = 0.04
(c)
K_{ph}^2 = 1.88, K_{DC}^2 = 3.69
D_0 = 200 Gy, σ_{ins} = 0.02
(d)
K_{ph}^2 = 1.88, K_{DC}^2 = 3.69
D_0 = 200 Gy, σ_{ins} = 0.04
D_e (Gy)
Natural dose (Gy)

Figure 8

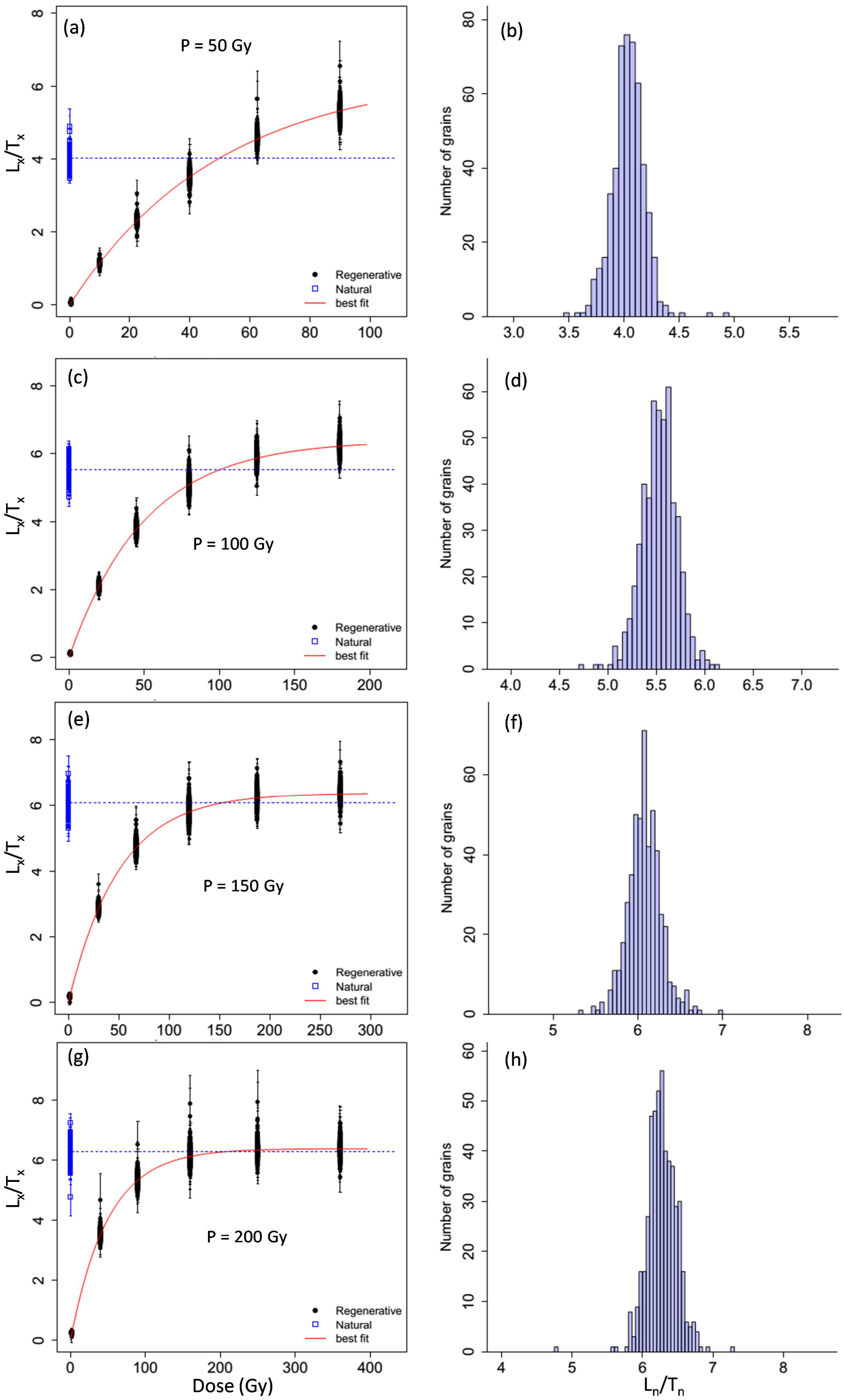

(a)
P = 50 Gy
Regenerative
Natural
best fit
(b)
(c)
P = 100 Gy
(d)
(e)
P = 150 Gy
(f)
(g)
P = 200 Gy
(h)
Dose (Gy)
Number of grains
L_x/T_x
L_n/T_n

Figure 9

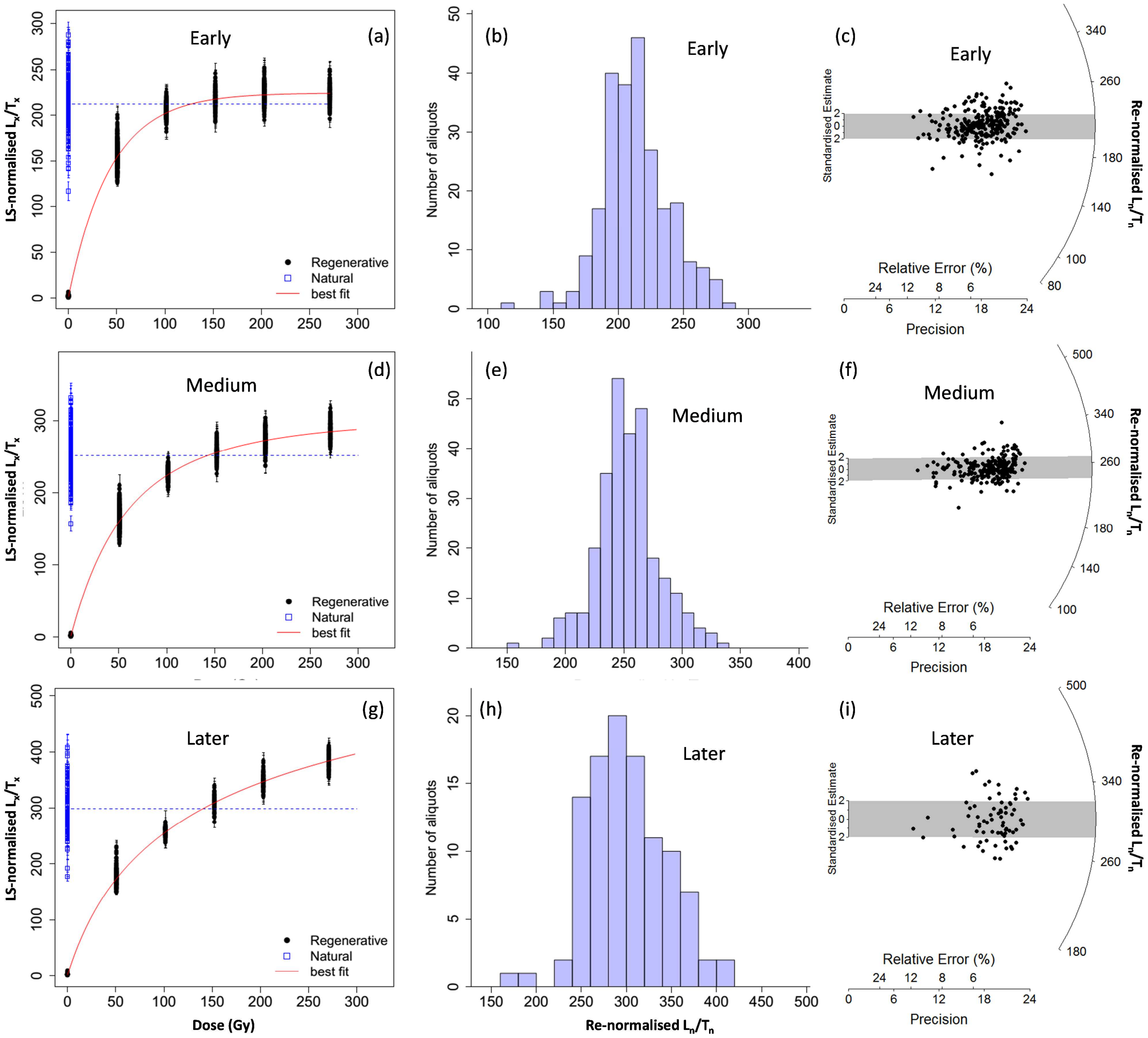

Early
(a)
LS-normalised L_x/T_x
Regenerative
Natural
best fit
(b)
Early
Number of aliquots
(c)
Early
Standardised Estimate
Relative Error (%)
Precision
Re-normalised L_n/T_n
Medium
(d)
(e)
Medium
(f)
Medium
Later
(g)
(h)
Later
(i)
Later
Dose (Gy)
Re-normalised L_n/T_n

Figure 10

Table 1: Summary of the mean D_0 values of simulated DRCs and the number of 'early saturated' grains for different groups of simulated grains with different natural doses, based on K_{ph}^2 and K_{DC}^2 values for three instruments and σ_{ins} values of 2% and 4%.

Instrument	K_{ph}^2	K_{DC}^2	**Instrumental uncertainty (σ_{ins})**	**True D_0 (Gy)**	**Simulated D_0 [a] (Gy)**	**Number of saturated grains [b]**			
						P = $2D_0$	**P = $3D_0$**	**P = $4D_0$**	**P = $5D_0$**
Risø2	1.88	3.69	2%	50	49.9 ± 3.0	3 (0.6%)	178 (35.6%)	350 (70.0%)	409 (81.8%)
				200	201.8 ± 20.9	7 (1.4%)	163 (32.6%)	328 (65.6%)	404 (80.8%)
			4%	50	50.0 ± 4.2	46 (9.2%)	292 (58.4%)	382 (76.4%)	424 (84.8%)
				200	203.6 ± 28.9	40 (8.0%)	269 (53.8%)	394 (78.8%)	429 (85.8%)
Ermintrude	1.23	4.49	2%	50	50.3 ± 3.0	7 (1.4%)	158 (31.6%)	341 (68.2%)	400 (80.0%)
				200	199.9 ± 16.7	5 (1.0%)	148 (29.6%)	330 (66.0%)	412 (82.4%)
			4%	50	50.3 ± 4.2	25 (5.0%)	280 (56.0%)	385 (77.0%)	427 (85.4%)
				200	202.7 ± 26.5	33 (6.6%)	275 (55.0%)	391 (78.2%)	420 (84%)
Moench	1.04	1.17	2%	50	49.9 ± 2.6	4 (0.8%)	138 (27.6%)	335 (67.0%)	402 (80.4%)
				200	201.0 ± 16.8	4 (0.8%)	140 (28.0%)	353 (70.6%)	426 (85.2%)
			4%	50	50.1 ± 4.0	34 (6.8%)	286 (57.2%)	409 (81.8%)	419 (83.8%)
				200	204.4 ± 26.6	33 (6.6%)	286 (57.2%)	389 (77.8%)	421 (84.2%)

[a] The simulated D_0 values are based on the mean of 500 simulated grains for each combination of K_{ph}^2 and K_{DC}^2 values and instrumental uncertainties. The uncertainty for each value represents one standard deviation.

[b] A total of 500 grains was simulated for each group. The percentage of saturated grains is shown in parantheses.