Ancient TL

www.ancienttl.org · ISSN: 2693-0935

Rodnight, H., 2008. *How many equivalent dose values are needed to obtain a reproducible distribution?* Ancient TL 26(1): 3-9. https://doi.org/10.26034/la.atl.2008.414

How many equivalent dose values are needed to obtain a reproducible distribution?

H. Rodnight[1,2]

1. Institute of Geography and Earth Sciences, Aberystwyth University, Ceredigion SY23 3DB, United Kingdom
2. Institut für Geologie und Paläeontologie, Universität Innsbruck, Innrain 52, 6020 Innsbruck, Austria (email: helena.rodnight@uibk.ac.at)

(Received 12 Dec 2007; in final form 16 May 2008)

Introduction

Since the development of single-aliquot measurement protocols, it has been feasible to rapidly obtain multiple estimates of the equivalent dose (D_e) for a single sample. Where a number of D_e values have been measured, a dose distribution can be obtained; the shape and spread of this distribution may be an important factor in obtaining the appropriate burial dose (D_b) for age calculation. Where the D_e values form a Gaussian distribution which is tightly clustered, some form of the mean is appropriate for the D_b value. Where the distribution is more scattered, however, a mean value is unlikely to be a good representation of the true burial dose. Scatter in D_e distributions can arise from a variety of factors including heterogeneous bleaching, post-depositional mixing and beta dose-rate heterogeneity; heterogeneous bleaching appears to be the most common cause of scatter and is most frequently discussed in the literature. Heterogeneous bleaching arises from insufficient exposure of a sediment to sunlight during transport, leading to residual trapped charge remaining in some or all of the grains on deposition; hence an overestimation of the burial dose is calculated for these grains. A relatively large number of studies have focussed on how to obtain an appropriate D_b value from such a distribution for a heterogeneously-bleached sample since the problem was first identified in a water-lain deposit by Murray et al. (1995). Less attention, however, has been paid to how to obtain a D_e dataset suitable for such analyses, i.e. how many replicate D_e measurements are sufficient to obtain a distribution that would result in a reproducible D_b value. If an insufficient number of D_e values is used, the final D_b value could be incorrect, regardless of how one calculated this D_b. Whilst for a well-bleached sample a relatively small number of D_e values are sufficient for D_b calculation, for a heterogeneously-bleached sample it might be expected that more measurements would be necessary to calculate the appropriate D_b.

The number of D_e values used in studies investigating partial bleaching varies considerably; whilst many studies obtain more than 50 values per sample (e.g. Olley et al., 1998; Lepper et al., 2000; Folz et al., 2001; Rowland et al., 2005) others use less than 10 D_e values for some samples (e.g. Colls et al., 2001; Fuchs and Lang, 2001; Srivastava et al., 2001). The quantity of material available for analysis can be a limiting factor in some instances; however, it is still desirable, where possible, to obtain enough D_e values for a reproducible distribution. The appropriate number of D_e values is not currently specified in the literature. This study aims to quantify this parameter through the use of a dataset obtained for a heterogeneously-bleached fluvial sample.

Sample details

Sample Aber/70KLA1, used in this study, was collected from an abandoned channel of the Klip River, South Africa, by augering through post-abandonment organic deposits until continuous sand was encountered. The uppermost part of this sandy deposit was sampled and interpreted as being derived from the final bedload transport event in the channel (Rodnight et al., 2006). The sample was pretreated following common procedures including 10% v.v. HCl and 20 volumes H_2O_2 to remove carbonates and organic matter, respectively. The 212-250 µm size fraction was obtained from dry sieving and sodium polytungstate solutions (densities of 2.62 and 2.70 g/cm^3) were used to obtain the quartz fraction which was etched with 40% HF acid for 45 minutes followed by washing with concentrated HCl. The sample was then resieved and the quartz grains retained were used for OSL measurements. Small aliquots (mask diameter 2 mm and containing ~ 30 grains) were used for all the measurements discussed in this paper. 175 aliquots of this sample were measured using the single-aliquot regenerative-dose (SAR) protocol (Murray and Wintle, 2000) with a preheat for 10 s at 220°C and a cut-heat at 160°C.

Aliquots were rejected if: (1) no detectable OSL signal was present after a regeneration dose had been applied; (2) the L_N/T_N value did not intersect with the growth curve; (3) the recycling ratio was not consistent with 1.0 ± 0.1; (4) the IR-OSL depletion ratio was not consistent with 1.0 ± 0.1 (Duller, 2003); (5) recuperation following a 0 Gy dose was detected (giving a L_X/T_X value greater than 5% of L_N/T_N). 122 D_e values were obtained after these rejection criteria had been applied. The errors associated with the individual D_e values were calculated in Luminescence Analyst (Version 3.20) from counting statistics, curve fitting and an instrumental reproducibility error of 2.5% (Duller, 2007).

The distribution obtained for Aber/70KLA1 indicated that the sample was heterogeneously bleached (Fig. 1), with a large range of D_e values present and an overdispersion value of 37% (Rodnight et al., 2006). Rodnight et al. (2006) demonstrated that the Finite Mixture Model (Galbraith and Green, 1990) gave the most reproducible D_b values when applied to replicate datasets of five samples and burial ages that were stratigraphically consistent with independent age control. Using the Finite Mixture Model on the dataset of 122 D_e values, a D_b of 2.38 ± 0.02 Gy was calculated for Aber/70KLA1. This was obtained by fitting 5 components to the data set and D_b was the value of the lowest dose component containing a minimum of 10% of the aliquots

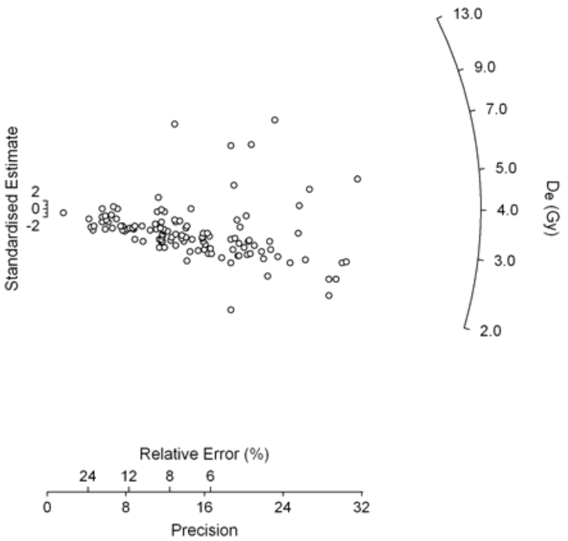


Figure 1: *Radial plot of D_e distribution for sample Aber/70KLA1. D_e values for 122 aliquots are shown. Redrawn from Rodnight et al. (2006).*

Methods

To determine how many D_e values are needed to characterise the distribution for Aber/70KLA1, sub-samples containing 5, 10, 15, 20, 30, 40, 50 and 60 D_e values were randomly selected from the 122, and this was repeated 20 times for each sub-sample size resulting in 160 datasets. When the sub-sample contains more than 5 D_e values there will be some replication of the D_e values in the datasets, but the analysis should still allow one to investigate the reproducibility of the distribution for a heterogeneously-bleached set of grains. Each of the 160 sub-samples (of 5, 10, 15, … 60 D_e values) was tested for normality using the 1-sample Kolmogorov-Smirnov test (SigmaPlot Version 7.0, SPSS Inc.). The results (Fig. 2) indicate that for this sample, at least 50 D_e values need to be measured to be certain of obtaining a dataset that is statistically non-normal. If the dataset contains less than 20 aliquots, there is a greater than 50% chance that a distribution from this heterogeneously-bleached sample will appear normal.

The D_b value was obtained for each of the 160 sub-samples using a variety of statistical models that have been proposed (Olley et al., 1998; Fuchs and Lang, 2001; Thomsen et al., 2003; Galbraith and Laslett, 1993; Galbraith and Green, 1990) for obtaining an appropriate D_b value from a heterogeneously-bleached sample. The change in the D_b values obtained, and the spread of the values for each sub-sample size can be used to assess how reproducible the results are for each method, at each sub-sample size. A brief description of each of these statistical techniques is given below.

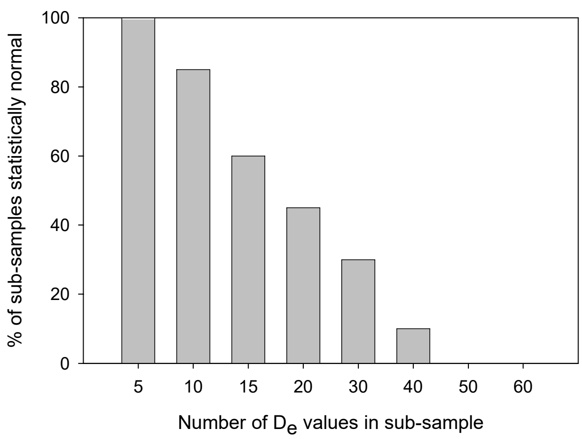


Figure 2: *Bar chart showing the percentage of the sub-sample datasets from sample Aber/70KLA1 that is statistically normal using the 1-sample Kolmogorov-Smirnov test.*

Method 1: Olley et al. (1998)
The authors of this paper found that they were able to calculate an OSL age consistent with the known age for a poorly bleached flood deposit from the Murrumbidgee River, Australia, by taking the mean of the lowest 5% of 78 measured D_e values. They subsequently used the mean of the lowest 5% of D_e values to calculate ages for fluvial samples from a core from the Namoi River; they found that the ages generally increased with depth. Whilst this method appeared to be suitable for the rivers investigated, it would be expected that the percentage of D_e values used would have to be 'calibrated' for different depositional systems.

Method 2: Fuchs and Lang (2001)
Low quantities of quartz were obtained from the samples detailed in this paper, therefore only nine or ten aliquots were analysed per sample. The results showed D_e values that were scattered more than expected from experimental variation, and this scatter was attributed to heterogeneous bleaching. Aliquots which had been artificially bleached and irradiated yielded a maximum standard deviation of 4% in the D_e values obtained from these measurements. To calculate a D_b value from the natural D_e values, based on only those D_e values from aliquots consisting of well-bleached grains, the D_e values for each sample were ranked in order from lowest to highest. Starting with the lowest value, and including one additional D_e value at a time, the mean D_e and the percentage standard deviation was calculated. This was repeated including aliquots with increasing values of D_e until the standard deviation of the mean was 4%, and this mean value was taken to be the most appropriate estimate of D_b.

Method 3: Thomsen et al. (2003)
This paper detailed a method which calculated the correct D_b from single grain D_e values for quartz extracted from irradiated blocks. To calculate a D_b value based on those grains from the well-bleached part of the distribution, the ratio of the external measurement of uncertainty (α_e) to the internal measurement of uncertainty (α_i) was used. The equations for these two measurements of uncertainty are:

$$\alpha_e^2 = \frac{\sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{\sigma_i^2}}{(n-1)\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}$$

$$\alpha_i^2 = \frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}$$

where x_i is the dose estimate from each individual grain, σ_i is its uncertainty, $\overline{x}$ is the weighted mean, and n is the number of measurements. α_e combines information on individual estimates of uncertainty for each grain and the deviation from a weighted mean. If there is no other source of error except for the uncertainty on the individual data points then α_e reduces to α_i, so for a large, normal population α_e/α_i tends to unity (i.e. where the overdispersion is 0). In a distribution containing partially bleached grains, this ratio can be used to determine which grains are well bleached, i.e. where the distribution in x_i is consistent with σ_i. Any additional variance because of heterogeneous bleaching will increase α_e relative to α_i. By ranking the individual equivalent doses from lowest to highest and calculating α_e/α_i for n = 2, 3…x until $\alpha_e/\alpha_i = 1 \pm (2(n-1))^{-0.5}$, only the well-bleached grains are used in the calculation of D_b. Any grain giving a D_e above this point is assumed to be partially bleached.

Method 4: Galbraith and Laslett (1993) – Minimum Age Model
The minimum age model was developed for samples where heterogeneous bleaching is evident (Galbraith et al., 1999). The model fits a truncated normal distribution to the logarithms of the individual D_e values, with the truncation point giving the value of D_b. To describe the distribution fitted to the dataset, four parameters and their errors need to be calculated: (1) the proportion of grains that were fully bleached prior to burial; (2) the truncation point of the distribution (i.e. the logarithm of the D_b value); (3) the value that would be the mean of the normal distribution fitted to the dataset; and (4) the overdispersion of this distribution.

Method 5: Galbraith and Green (1990) – Finite Mixture Model
This model was developed for instances where grains of more than one discrete population are present and where each population has been well bleached and can be described by the central age model (Galbraith et al., 1999). Although this model has been designed for a distribution consisting of a discrete number of populations (components), it can be applied to a heterogeneously-bleached sample. As the model selects populations based on the logarithms of D_e values that are consistent with one another (within a pre-defined overdispersion value), the D_b of the lowest population will be essentially derived from a

normal distribution of the lowest D_e values. Thus, if the dataset contains some D_e values measured from well-bleached grains, this lowest population should give the appropriate D_b value for the sample. To analyse a dataset, the model is run repeatedly, starting with only one component, and then with an additional component included each time. The model output includes two parameters which can be used to select the most appropriate number of components for fitting the dataset. The maximum log likelihood (llik) of a fit is likely to improve indefinitely as a greater number of components are fitted, although this does not necessarily mean that the solution is a better one. The Bayesian Information Criterion (BIC) takes into account the complexity of the model as well as the goodness of the fit to the data, and hence reduces down to a minimum at the 'best' fit, before rising as the increase in the llik is outweighed by the additional components (Jacobs et al., 2008). For each of the 160 sub-samples in this study, the D_b value was calculated from running the model with the number of components that were calculated to have the best-fit (using the BIC). The final D_b value calculated for each sub-sample was based on the lowest population that contained at least 10% of the D_e values in the dataset. This 10% value was selected arbitrarily so that populations based on one or two D_e values only were not used for derivation of the sample D_b. The number of components found to have the best fit varied between the sub-samples. As an example, the number of components fitted to the 50 D_e sub-samples was as follows: two components (2 sub-samples); three components (6 sub-samples); four components (8 sub-samples); five components (3 sub-samples); and six components (1 sub-sample).

The first two methods do not incorporate the error on the individual D_e values in the analysis. If the brightness of the OSL signal is variable on an aliquot-to-aliquot basis, however, then the D_e errors will also vary considerably. Thus it is preferable for them be taken into account, otherwise a value that may appear to be an outlier can actually form part of the main distribution within errors.

For Method 5 an overdispersion value of 10% was used for each component; since no well-bleached samples were available from the Klip River study area this was based on a dataset of D_e values obtained from analysis of small aliquots of a well-bleached last glacial maximum linear sand dune from Tasmania (see Rodnight et al., 2006 for further details). For Method 4 this 10% overdispersion was incorporated into the minimum age model so that the results were comparable with those from Method 5.

Results

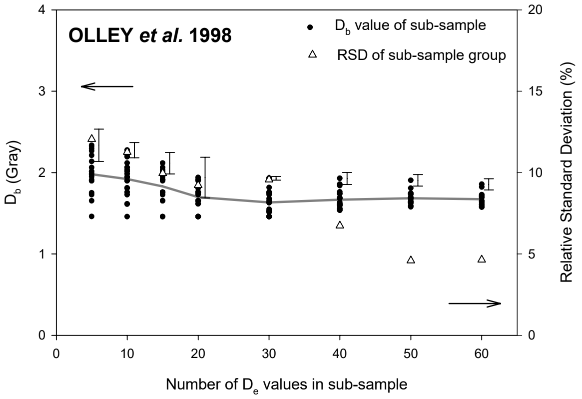
The results obtained for each method using the dataset for Aber/70KLA1 are shown in Figure 3. The D_b value calculated for each of the 160 datasets is plotted as a filled circle as a function of sub-sample size. The error bars indicate the absolute error associated with the greatest D_b value calculated for each sub-sample size. The grey line joins the mean D_b value for each sub-sample size to indicate whether changing sub-sample size can be associated with the trend in the mean D_b value calculated. The relative standard deviation (RSD) calculated for the 20 D_b values for each sub-sample size is indicated by the open triangle.

The results using Method 1, of Olley et al. (1998), to calculate D_b are shown in Figure 3a. When the sub-sample consists of more than 20 D_e values, consistent values are obtained for the D_b. The D_b for datasets with $\leq$20 D_e values will be derived from the lowest D_e value only, whilst the D_b for the 60 D_e sub-sample is based on the lowest 3 D_e values. As this technique takes into account only very few aliquots from the lower end of the distribution, each D_b value is essentially based on the same few D_e values. This is demonstrated by the mean D_b values that fall monotonically for between 5 and 20 D_e values, and then remain constant at around 1.8 Gy. Using this method, the RSD of the results falls to <5% when at least 50 D_e values are included in the analysis.

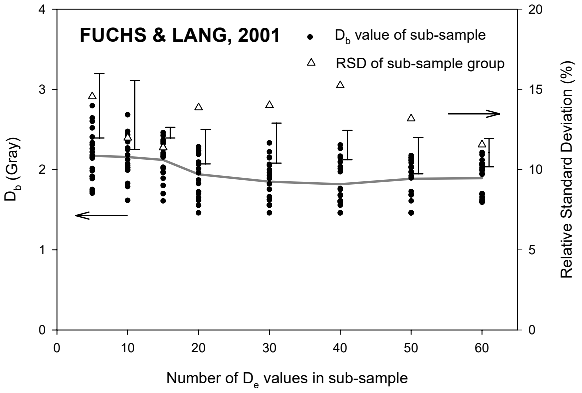
The RSD of the D_b values obtained using Method 2 (Fuchs and Lang, 2001) show poor reproducibility in the results (Fig. 3b). This poor reproducibility arises because of the sensitivity of the procedure to a high percentage standard deviation in the lowest two or three D_e values in the sub-sample. For instance, if the first few D_e values vary considerably, a high relative standard deviation results and the D_b is calculated from only these very few values. Even when the sub-sample consists of 60 D_e values, the RSD of the D_b values remains high (>10%).

For sub-samples containing $\leq$15 D_e values, Method 3 (Thomsen et al., 2003) tends to select low outlying values for calculation of the D_b; however, with larger sub-sample sizes this does not occur (Fig. 3c). In general, the mean D_b value that is calculated stays relatively consistent at ~2.1 Gy, and the RSD of the D_b values is <5% when 50 D_e values are included in the analysis. The errors associated with the D_b value are, however, highly variable owing to the manner in which the α_e/α_i ratio is used to derive the errors.

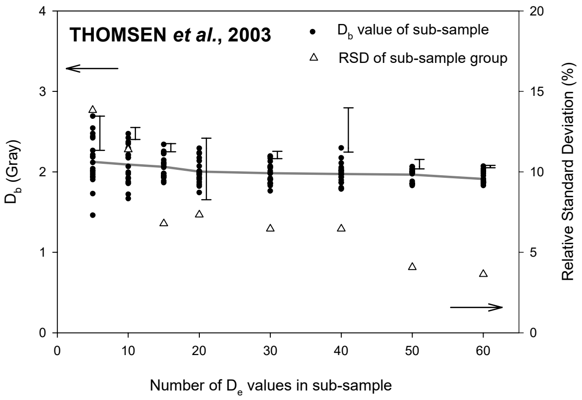
a)



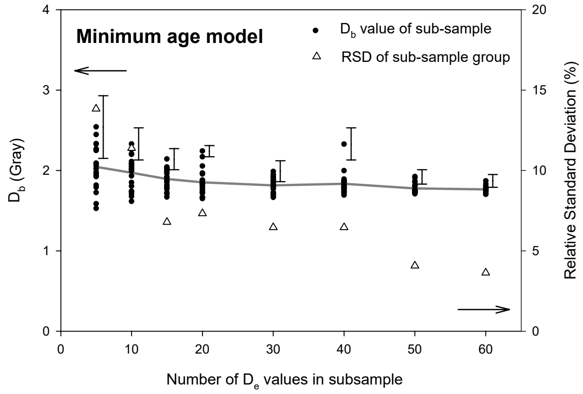
b)



c)



d)



e)

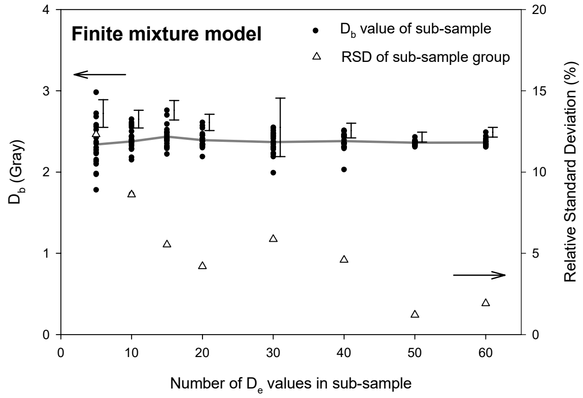


Figure 3: *Results from analysis of different sub-samples of the D_e dataset. The sub-samples were used to derive D_b values (left-hand axis) using five statistical methods, the D_b for each sub-sample is shown as a filled circle; the errors associated with the largest D_b value for each sub-sample size are also shown. The grey line shows the mean D_b values for each sub-sample size. The relative standard deviations of the D_b values for each size grouping are plotted as triangles using the right-hand axis.*

Method	Minimum D_b (Gy)	Maximum D_b (Gy)	Mean D_b (Gy)
1. Olley et al. (1998)	1.58 ± 0.08	1.90 ± 0.01	1.68 ± 0.02
2. Fuchs and Lang (2001)	1.46 ± 0.32	2.17 ± 0.23	1.88 ± 0.06
3. Thomsen et al. (2003)	1.83 ± 0.28	2.07 ± 0.06	1.97 ± 0.02
4. Minimum Age Model	1.71 ± 0.10	1.92 ± 0.09	1.78 ± 0.01
5. Finite Mixture Model	2.31 ± 0.05	2.43 ± 0.05	2.36 ± 0.01

Table 1: *Summary of results obtained using the sub-samples containing 50 D_e values. Showing the minimum and maximum D_b values, and the mean D_b (and its standard error) of the 20 calculated values.*

Using Method 4, the Minimum Age Model of Galbraith and Laslett (1993), for D_b analysis the average D_b values decrease until a sub-sample size of 30 after which they remain relatively constant (Fig. 3d). A sub-sample size of 50 D_e values or more is needed to reduce the RSD to less than 5%.

Using the Finite Mixture Model (Method 5), of Galbraith and Green (1990), the results (Fig. 3e) are similar to the Minimum Age Model; the mean D_b values fluctuate until a sub-sample size of 30 is used, after which they remain relatively constant, and a sub-sample size of 50 D_e values or more is needed to reduce the RSD to less than 5%. As one would expect, however, the minimum age model consistently derives lower D_b values than the finite mixture model.

The minimum and maximum D_b values calculated from the 20 sub-samples containing 50 D_e values are detailed in Table 1. This shows that Methods 1 and 4 (Olley et al., 1998, and the Minimum Age Model) calculate the lowest D_b values for all the datasets, whilst the Finite Mixture Model calculated the highest D_b values. The errors associated with the D_b values vary considerable as a result of the different ways in which they are calculated for the individual methods. Between the various methods there is up to 40% difference in the mean D_b values, demonstrating the necessity of careful consideration when choosing a statistical technique for D_b calculation.

Discussion and Conclusions

With the exception of Method 2, the RSD of the D_b values calculated from sub-samples is always less than 5% for sub-samples containing 50 D_e values. In general, the results from Methods 2 and 3 show more variation in the final D_b values than the other models; this suggests that techniques based on the inclusion of increasing values until a predefined parameter is reached encounter problems with low, outlying, values. This is particularly demonstrated by the results for Method 2 owing to the fact that the errors on the individual D_e values are not taken into consideration. Only the D_b values calculated using the Finite Mixture Model are consistent with the result obtained from the entire dataset of 122 D_e values (2.38 ± 0.02 Gy, calculated using the Finite Mixture Model) as one might expect.

Sample Aber/70KLA1 has an overdispersion parameter of 37%, other work has detailed D_e distributions with similar or greater values of overdispersion (e.g. Olley et al., 2004; Arnold et al., 2007), demonstrating that it is not an unusual situation. The results from the normality tests on the sub-samples of Aber/70KLA1 suggest that 50 D_e values are necessary to be certain of determining whether the sample is non-normal; datasets consisting of less than 50 D_e values may appear to be well-bleached when in fact the sample is not. The results from the different statistical models also indicate that, for most of the methods, at least 50 D_e values are necessary to obtain reproducible estimates of D_b for this sample, and the D_b values calculated using the Finite Mixture Model were consistent with the D_b from the entire dataset of 122 D_e values. Therefore, for analysis of samples from a similar depositional environment which may be heterogeneously bleached, 50 D_e values are recommended as the minimum working population.

References

Arnold, L.J., Bailey, R.M., Tucker, G.E. (2007). Statistical treatment of fluvial dose distributions from southern Colorado arroyo deposits. *Quaternary Geochronology* **2**, 162-167.

Colls, A.E., Stokes, S., Blum, M.D., Straffin, E. (2001). Age limits on the late Quaternary evolution of the upper Loire River. *Quaternary Science Reviews* **20**, 743-450,

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

Duller, G.A.T. (2007). Assessing the error on equivalent dose estimates derived from single aliquot regenerative dose measurements. *Ancient TL* **25**, 15-24.

Folz, E., Bodu, P., Bonte, P., Joron, J.L., Mercier, N., Reyss, J.L. (2001). OSL dating of fluvial quartz from Le Closeau, a Late Paleolithic site near Paris – comparison with 14C chronology. *Quaternary Science Reviews* **20**, 927-933.

Fuchs, M., Lang, A. (2001). OSL dating of coarse-grain quartz using single-aliquot protocols on sediments from NE Peloponnese, Greece. *Quaternary Science Reviews* **20**, 783-787.

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., Laslett, G. (1993). Statistical models for mixed fission track ages. *Radiation Measurements* **21**, 459-470.

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.

Jacobs, Z., Wintle, A.G., Duller, G.A.T., Roberts, R.G., Wadley, L. (2008). New ages for the post-Howiesons Poort, late and final Middle Stone Age at Sibudu, South Africa. *Journal of Archaeological Science* **35**, 1790-1807.

Lepper, K., Agersnap Larsen, N., McKeever, S.W.S. (2000). Equivalent dose distribution analysis of Holocene eolian and fluvial quartz sands from Central Oklahoma. *Radiation Measurements* **32**, 603-608.

Murray, A.S., Wintle, A.G. (2000). Luminescence dating of quartz using an improved single-aliquot regenerative-dose protocol. *Radiation Measurements* **32**, 57-73.

Murray, A.S., Olley, J.M., Caitcheon, G.G. (1995). Measurement of equivalent doses in quartz from contemporary water-lain sediments using optically stimulated luminescence. *Quaternary Geochronology (QSR)* **14**, 365-371.

Olley, J.M., Caitcheon, G.G., Murray, A.S. (1998). The distribution of apparent dose as determined by optically stimulated luminescence in small aliquots of fluvial quartz: implications for dating young sediments. *Quaternary Geochronology (QSR)* **17**, 1033-1040.

Olley, J.M., Pietsch, T., Roberts, R.G. (2004). Optical dating of Holocene sediments from a variety of geomorphic settings using single grains of quartz. *Geomorphology* **60**, 337-358.

Rodnight, H., Duller, G.A.T., Wintle, A. G., Tooth, S. (2006). Assessing the reproducibility and accuracy of optical dating of fluvial deposits. *Quaternary Geochronology* **1**, 109-120.

Rowland, J.C., Lepper, K., Dietrich, W.E., Wilson, C.J., Sheldon, R. (2005). Tie channel sedimentation rates, oxbow formation age and channel migration rate from optically stimulated luminescence (OSL) analysis of floodplain deposits. *Earth Surface Processes and Landforms* **30**, 1161-1179.

Srivastava, P., Juyal, N., Singhvi, A.K., Wasson, R.J., Bateman, M.D. (2001). Luminescence chronology of river adjustment and incision of Quaternary sediments in the alluvial plain of the Sabarmati River, north Gujarat, India. *Geomorphology* **36**, 217-229.

Thomsen, K.J., Jain, M., Bøtter-Jensen, L., Murray, A.S., Jungner, H. (2003). Variation with depth of dose distributions in single grains of quartz extracted from an irradiated concrete block. *Radiation Measurements* **37**, 315-321.

Reviewer

A. Lang