Estimating the error in equivalent dose values obtained from SAR

G. W. Berger

Desert Research Institute, 2215 Raggio Parkway, Reno, NV 89512, USA
(e-mail: glenn.berger@dri.edu)

(Received 26 September 2010; Final form 4 December 2010)

Abstract
Central to all aspects of application of SAR (Single-Aliquot-Regenerative-dose) procedures to dating is the accurate estimation of errors in equivalent dose values (D_e). Some approximation approaches have been outlined by others. Here a more rigorous approach is outlined that incorporates the contributions to the total error in D_e from not only error in L_0/T_0 (undosed and unbleached ratio) and the contribution from the scatter of data about the best-fit curve, but also contributions from errors in the best-fit regression parameters and from their covariances. Unequal weighting of data points is included. Algorithms and procedures are detailed for two [saturating exponential (E), and saturating exponential-plus-linear (E+L)] of the 4 most common dose-response models [which include linear (L) and quadratic (Q)]. Computed D_e values and errors are compared for 6 new and three previously published data sets, using the approach of this paper and two of the approaches of Duller (2007). Differences in D_e and its error estimate from these comparisons are often, but not always, small and insignificant, thus indicating that errors in best-fit parameters and their covariances make a small contribution to total error in D_e in most situations. One example data set evaluates the effects of dose-point spacing on the estimated D_e error.

Introduction
As in Duller (2007), Berger (1990, hereinafter "B90") and Berger et al. (1987, hereinafter, "BLK87"), and some other authors (e.g., Grün and Rhodes, 1992), the error analyses presented herein are intended to capture the effects of random errors, not systematic errors. Discussions about the relative role of random and systematic errors in the construction of dose-response curves (DRCs) for TL (thermo-luminescence), ESR (electron spin resonance), and OSL (optically stimulated luminescence) data have been joined by many authors (e.g., BLK87; Grün and Rhodes, 1992; Grün and Packman, 1993; Hayes et al., 1998; Galbraith, 2002). In the case of SAR applications, low signals could lead to dominance of systematic errors whereas larger signals are expected to lead to dominance of random errors in L/T ratios (test-dose-normalized OSL in the SAR approach).

The widely used *Analyst* software package that accompanies the Risø luminescence reader systems provides SAR D_e regression procedures for the L, Q, E and E+L DRC models.

The D_e error estimation procedures of *Analyst* are summarized by Duller (2007). His procedure essentially combines in quadrature two components of the total analytical error in D_e: an interpolated-range (L_0/T_0 ±error) effect, and the effect of scatter (Duller's equation 7) of L/T ratios about the best-fit curve. This interpolated-range procedure is adopted here in the E+L regression model. Duller's (2007) approach does not capture the effects of errors in the fitting parameters, nor any effects of covariances among the errors in these parameters. His approach does, however (Duller, pers. comm., 2010) incorporate the effects of assigning unequal weights to the data points, where weighting is by inverse variance of the estimated absolute errors in L/T ratios.

The importance of weighting in regression and error analysis has been emphasized and discussed elsewhere (e.g., BLK87; B90; Grün and Rhodes, 1992; Grün and Packman, 1993; Hayes et al., 1998) with regard to constructing DRCs for extrapolation to D_e values (in TL and ESR dating). For such extrapolated regressions, Hayes et al. (1998) point out that Bluszcz (1988) discussed how to deal (using the Monte Carlo approach) with TL/ESR cases of mixed error types (e.g., combinations of constant absolute errors, variable absolute errors, constant relative errors). The above cited authors point out (implicitly or explicitly) that dominance by systematic errors implies constant absolute errors, while dominance by random errors can imply constant relative errors. Grün and Rhodes (1992)

conclude that the effects of constant relative errors can be captured by use of weighting by inverse variance, and that this choice leads to smaller errors in estimated D_e values from extrapolated saturating-exponential (hereinafter, 'E fit') TL/ESR DRCs than does the use of equal weighting, which captures the effects of constant absolute errors. This conclusion of Grün and Rhodes (1992) supports the choice of BLK87 for the use of weighting by inverse variance in the calculation of D_e errors from extrapolated DRCs.

While the above efforts concerned extrapolation of DRCs, SAR employs interpolation (as does the regeneration TL procedure). Therefore, estimation of errors in SAR D_e values may have a different dependency on the germane empirical factors than in the case of extrapolations. With SAR, because of the nature of the measurement of L/T ratios, estimation of variance in each SAR dose-data point is quite tractable (e.g., Galbraith, 2002). Thus use of weighting by inverse variance of L/T ratios as presented in this paper is straightforward, avoiding the kinds of uncertainties embedded with the measurement of (for example) TL signals (e.g., BLK87, Appendix A).

Here the mathematical procedures of BLK87 and B90 are adapted to the SAR conditions. BLK87 and B90 used the Gauss-Newton method of linearization of the mathematically non-linear E and E+L models to obtain fitting parameters, and the 'delta' method to obtain error estimates in the extrapolated D_e values. The procedures of BLK87 are statistically accurate when relative errors in the data points are small (e.g., <5%).

Results from these adapted procedures are compared below to the results from the two error-estimation approaches of Duller (2007) as executed in *Analyst* 2007 (v.3.24): his 'curve-fitting' and Monte Carlo approaches. His 'curve-fitting' approach captures the effects mentioned above (error in L_0/T_0, unequal weighting, data scatter about the DRC). His Monte Carlo approach presumably captures the end effects of all empirical variables. As Duller (2007) reminds us, when L_0/T_0 is near the saturation value of E fits, the error in D_e, likely will not be symmetrical. In this situation, only the Monte Carlo approach may yield a statistically accurate estimate of the error in D_e. To construct the best-fit DRCs, Duller (2007) employed the Levenberg-Marquardt (L-M) algorithm. This has many advantages (e.g., Press et al., 1986; Hayes et al., 1998), not the least of which can be speed of convergence for mathematically non-linear models. Essentially then, the algorithms of this paper employ weighting of L/T ratios by inverse variance of same, and, moving beyond the approach of Duller (2007), capture the effects of errors and covariances in fitting parameters.

Although the models presented below force the dose-response curve through the origin, they could be modified to permit passing the regression through the recuperation datum. The 'curve-fitting' procedure of Duller (2007) passes the regression curve through the recuperation datum, but his *Analyst* software permits the choice of forcing the curve through the origin. Since one of the data-acceptance criteria (e.g., Wintle and Murray, 2006) in the use of SAR is the rejection of data for which the recuperation is $>1\sigma$ or $>2\sigma$ from zero, forcing the curve through the origin is likely to be usefully accurate in almost all SAR situations. Furthermore, it has been demonstrated empirically (e.g., Ballarini et al., 2007; Berger, 2009; Berger et al., 2010; Cunningham and Wallinga, 2010) that placing the interval selected for 'background' subtraction close to the beginning part of the OSL decay curve, rather than at the end, often has the effect of reducing the recuperation signal to near zero.

Some Nomenclature

The following terms are employed: σ^2 = absolute-error variance; N = number of L/T (hereinafter, "y") data points, including the origin; w = weight, which for SAR is $1/\sigma_y^2$; $y_0 = L_0/T_0$; θ = a fitting parameter (e.g., a, b, or c below); scalar VAR = weighted sum of squares of residuals (an estimate of the scatter of data about the best-fit curve). Note that σ_y^2 can be calculated appropriately by the method of Galbraith (2002). Symmetrical matrix SIG is the matrix Ψ of equation 10 in BLK87 and of equation 4 in B90; and

$$SIG = VAR \cdot (I)^{-1} \qquad (1)$$

That is, SIG is the product of scalar VAR and the inverse matrix of I. Matrix I is the information matrix of BLK87, and thus SIG is the variance-covariance matrix, or error matrix. The diagonal elements of SIG give the variances in the individual fitting parameters, while the off-diagonal elements give the covariances.

Employed throughout this paper is the standard propagation-of-variance equation

$$\sigma^2 = \sum_{k,s} \frac{\partial f}{\partial \theta_k} \frac{\partial f}{\partial \theta_s} \sigma_{\theta_k} \sigma_{\theta_s} \qquad (2)$$

where f is the model function for the curve, and the paired terms provide the variances when k = s and the covariances when $k \neq s$. The covariance terms can be

ignored (set to zero) only if the errors in the fitting parameters are independently distributed (independent of each other) and symmetrical with respect to positive and negative values. We assume here that these errors are Gaussian (although this assumption is not required in the approach of BLK87), hence symmetrical, but we cannot assume independence of errors in fitting parameters. However, it is assumed realistically that the error in y_0 is independent of the errors in the fitting parameters. Note that hereinafter, the subscript "i" will be used for summations over the non-zero dose-axis (x) data points.

The elements of the matrix I (when needed below) are derived from equation 11 of BLK87, which reduces to

$$I_{k,s} = \sum_i \frac{1}{f_i^2} \frac{\partial(f_i)}{\partial\theta_k} \cdot \frac{\partial(f_i)}{\partial\theta_s} \qquad (3)$$

Linear Fit

This model, a linear dose response curve passing through the origin, applies to many young sediments or heated materials. Realistically we assume that there are errors in y, not in x. Using the weighted least-squares principle, we wish to minimize

$$S = \sum_i w_i (y_i - f_i)^2 \qquad (4)$$

where f = bx. Deming's (1964) equation 39 (p. 33) provides a straightforward form of S that circumvents the need to know b beforehand, and his equation 34 (p.31) permits easy calculation of slope b.

Inclusion of the effects of scatter of data about the best-fit curve is required. Deming's (1964) equation 41 (p.34) provides a simple formulation of σ_b^2 that includes the data-scatter effect, which he terms the 'external standard error' in b. In our nomenclature, his divisor m-1 in this equation is N-2. Thus, the total variance in D_e is

$$\sigma_{D_E}^2 = D_E^2 \left[\left(\frac{\sigma_b}{b}\right)^2 + \left(\frac{\sigma_{y_0}}{y_0}\right)^2 \right] \qquad (5)$$

Here we assume realistically that there is no covariance of y_0 and b, that is, that the errors in y_0 and b are independent.

Quadratic Fit

This model applies to young sediments for which the dose response curve is supralinear (e.g., Berger, 1987) or to dose response curves that are slightly sublinear. As BLK87 show, in the low dose region of a dose response, a second-order polynomial can be a good approximation to the physically realistic saturating exponential. The Q model is

$$f = bx + cx^2 \qquad (6)$$

and we wish to minimize equation 4 to obtain estimates of b and c.

We use the matrix normal equations, simplified to the case of only one curve (not the intersection of two curves as in BLK87). Then in the case of unequal weights, the coefficients b and c are calculated from equation 2 of BLK87. Following the procedural steps of BLK87, we iterate the calculation of b and c until the difference between successive estimates meets some pre-set convergence limit. The D_e value is the absolute value of the solution to equation 6 when f is replaced by y_0 and x is replaced by D_e.

To obtain the errors in b and c, we need scalar VAR and matrix I from equation 1. Then

$$VAR = \frac{\sum_i w_i (y_i - f_i)^2}{N-3} \qquad (7)$$

with denominator N-3 because there are only two fitting parameters. From equation 3, the elements of matrix I are $I_{bb} = \sum_i w_i x_i^2$, $I_{cc} = \sum_i w_i x_i^4$, $I_{bc} = I_{cb} = \sum_i w_i x_i^3$. Here, the covariances (off-diagonal elements) are not zero because the errors in b and c are not independently distributed. Therefore, using equations 1 and 7, we obtain the elements SIG_{bb}, SIG_{cc} and SIG_{bc} of the error matrix SIG.

The total variance in D_e follows from equation 2 as

$$\sigma_{D_E}^2 = \left(\frac{\partial f}{\partial y_o}\right)^2 \sigma_{y_0}^2 + \left(\frac{\partial f}{\partial b}\right)^2 SIG_{bb} + \left(\frac{\partial f}{\partial c}\right)^2 SIG_{cc} + 2\left[\frac{\partial f}{\partial b} \cdot \frac{\partial f}{\partial c} \cdot SIG_{bc}\right] \qquad (8)$$

where we assume reasonably that the error in y_0 is independent of the errors in b and c. Therefore, applying equation 8 to the analytic formula for D_e (solution to equation 6), equation 8 becomes

$$\sigma_{D_E}^2 = \left(\frac{\sigma_{y_0}}{Z}\right)^2 + \left(\frac{b}{2c(Z-1)}\right)^2 SIG_{bb} + \left(\frac{y_0 - ZD_E}{cZ}\right)^2 SIG_{cc} + 2\left[\left(\frac{b}{2c(Z-1)}\right)\left(\frac{y_0 - ZD_E}{cZ}\right) SIG_{bc}\right] \qquad (9)$$

where $Z = |b^2 - 4c(-y_0)|$.

Saturating Exponential (E)
This model has the most general applicability to charge-trap physical processes, for which trap-type filling at high doses can occur. The model is

$$f = a(1-e^{-bx}) \quad (10)$$

and we wish to minimize equation 4 to obtain estimates of parameters a and b. From equation 10,

$$D_e = \frac{1}{b} ln \left[\frac{a}{a-y_0}\right] \quad (11)$$

since $f = y_0$ when $x = D_e$.

Regression to obtain a and b
We use the matrix methods and follow the steps of BLK87. The suggested procedure of BLK87 begins with obtaining initial (trial) estimates of a and b by setting the weight matrix = 1 and using a quadratic form of the dose-response curve. This step gives initial estimates of a and b.

Then using estimated (calculated) weights, we find increments to a and b using equation 9 of BLK87. In practice, equation 9 can be reformulated (BLK87 matrix symbol β replaced here by matrix symbol ΔA) to:

$$\Delta A = ([WU]^t[WU])^{-1}([WU]^t[WY^*]) \quad (12)$$

where the matrix elements are:
$wu_a = (1 - e^{-bx_i})\sqrt{w_i}$,
$wu_b = ax_i e^{-bx_i}\sqrt{w_i}$,
$wy^* = y_i - a(1 - e^{-bx_i})\sqrt{w_i}$.

Equation 12 is iterated to the desired level of convergence of a and b. Then D_e is calculated from equation 11.

Error in D_e
We use equation 2, with the D_e from equation 11 replacing f in equation 2. Thus the variance in D_e is calculated from equation 8 (with b, c replaced by a, b respectively). Using equation 10, we obtain the elements of I:
$I_{aa} = \sum_i w_i$,
$I_{bb} = \sum_i w_i (x_i e^{-bx_i})^2$,
$I_{ab} = I_{ba} = \sum_i w_i f_i e^{-bx_i}$

Elements of SIG are obtained from equation 1. Using equation 11 for D_e, the partial derivatives in equation 8 become:

$$\frac{\partial D_E}{\partial y_0} = [b(a - y_0]^{-1},$$
$$\frac{\partial D_E}{\partial a} = -y_0/ab(a - y_0) \text{ and}$$
$$\frac{\partial D_E}{\partial b} = -D_E/b .$$

These results then permit the calculation of the variance in D_e.

Saturating Exponential Plus Linear (E+L)
This model, as stated by B90, is physically realistic when charge-trap creation occurs at high doses. The linear term can represent the trap-type creation region of the dose response curve. Dose response curves with apparent linear growth superimposed upon a saturating exponential growth have been observed under laboratory irradiations in TL dating (e.g., B90) and quartz SAR OSL dating (e.g., Murray et al., 2008; Pawley et al., 2008). Does this E+L laboratory response of test-dose-normalized OSL require the assumption of E+L trap filling under natural irradiation (sediment burial) conditions? Notwithstanding this question, comparison of quartz SAR OSL age estimates derived from use of an E+L dose response curve model with independent ages have shown good agreement at ~ 200 ka (Seyda River sites, Murray et al., 2008) and at ~ 450 ka (Pawley et al., 2008). On the other hand, Lai (2010) reported dramatic age under-estimation for old loess when E+L dose response curves were observed. Could this under-estimation be partly an effect of differences in his SAR procedure from those of others (e.g., Murray and Wintle, 2003)?

There is another physically realistic model that can mimic an apparent E+L dose response curve in SAR, and this is the 'double-saturating-exponential' (DSE) model (e.g., Wintle and Murray, 2006). Here the apparent linear term can manifest a second set of charge traps with a different (larger) saturation level than the first type. Murray et al. (2007) employed this model with SAR data for quartz from their Sula River sites, but obtained age underestimates. Murray et al. (2008) discussed this result, and suggested that perhaps the nature of the geological setting at Sula River and of the independent age assignments might have had more influence on this underestimate than any putative instability in the quartz SAR signal from the high-dose region of the dose response curve. However, in this report only the E+L model is considered for error analysis because this model is much more tractable than the DSE model.
The model is

$$f = a(1-e^{-bx}) + cx \quad (13)$$

A straightforward, efficient and acceptably accurate path to follow is to use equation 13 to obtain a, b, c, VAR and SIG, then to use

$$f' = y_0 - a(1-e^{-bx}) - cx \qquad (14)$$

to solve for D_e, because $f' = 0$ when $x = D_e$. Equation 14 can be solved for D_e very efficiently by using the Newton-Raphson iterative procedure (e.g., McCalla, 1967).

Regression to obtain a, b, c
As for the E case, we use the quadratic model with weights = 1 to obtain initial estimates of a and b. Then we assign an initial estimate for c of (e.g.) 0.005, based on the likely general increase in L/T as a fraction of applied dose in units of sec. With these initial estimates of a, b and c, we use equation 12 to begin the iterative process to refine these estimates toward some limit of convergence. In this case, the elements of matrices WU and WY* become:

$wu_a = (1 - e^{-bx_i})\sqrt{w_i}$,
$wu_b = (ax_i e^{-bx_i})\sqrt{w_i}$,
$wu_c = x_i\sqrt{w_i}$, and
$wy^* = [y_i - a(1 - e^{-bx_i}) - cx_i]\sqrt{w_i}$.

Solution for D_e and error in D_e
Once we have a, b and c, we can use equations 14 to iteratively solve for D_e. Here, the estimation of errors in D_e is more complex than for the previous models. An efficient, fast and reasonably accurate procedure is to approximate the total variance in D_e as

$$\sigma^2_{D_E} = \sigma^2_{D_E}(y_0) + \sigma^2_{D_E}(\theta) \qquad (15)$$

which is the sum of variances from the effect on D_e of error in y_0 and the variance arising from the combination of the effects of data scatter about the best-fit curve with the variances and covariances in the fitting parameters a, b and c.

Error in D_e arising from error in y_0
An historically effective way (and used by Duller, 2007) to estimate the component of variance in D_e that relates to the effects of the variance in y_0 is to iterate the solution of equation 14 under 3 conditions:

(i) set $y_0' = y_0 + \sigma_{y_0}$, obtain D_e' ;
(ii) set $y_0'' = y_0 - \sigma_{y_0}$, obtain D_e'' ;
and (iii) set $y_0' = y_0$, obtain the best-fit D_e.

Although $\sigma^2_{D_E}(y_0)$ may not always be symmetrical about D_e, it is acceptable to set

$$\sigma^2_{D_E}(y_0) = [(D_E' - D_E'')/2]^2 \qquad (16)$$

Error in D_e arising from data scatter and from errors in a, b and c
For this purpose we can use the efficient equation 4 of B90 (with his Ψ replaced by our SIG), which captures all the effects we wish to incorporate. Then in that equation 4,

$V^t = \left(\frac{\partial f}{\partial a}, \frac{\partial f}{\partial b}, \frac{\partial f}{\partial c}\right)$
$= ([1 - e^{-bx}],\ axe^{-bx},\ x)$, and
$\frac{\partial f}{\partial D_E} = c + abe^{-bD_E}$.

For use in equation 1 we need
$VAR = \sum_i \frac{w_i(y_i - f_i)^2}{N-4}$

where f_i is given by equation 13. The elements of I, derived using equation 3, are:

$I_{aa} = \sum_i w_i(1 - e^{-bx_i})^2$,
$I_{bb} = \sum_i w_i(ax_i e^{-bx_i})^2$,
$I_{cc} = \sum_i w_i {x_i}^2$,
$I_{ab} = I_{ba} = \sum_i w_i a x_i e^{-bx_i}(1 - e^{-bx_i})$,
$I_{ac} = I_{ca} = \sum_i w_i x_i(1 - e^{-bx_i})$,
$I_{bc} = I_{cb} = \sum_i w_i a x_i^2 e^{-bx_i}$.

Therefore, with these equations for the elements of matrix I and the equation for VAR, we obtain $\sigma^2_{D_E}(\theta)$ from equation 4 of B90 and the total variance in D_e from equation 15 combined with equation 16.

Comparison of Results from Data Sets
Results from 3 error-estimation approaches are compared below. The first two approaches are the 'curve-fitting' and Monte Carlo routines of Duller (2007) outlined in the first section and executed in *Analyst* 2007, and the third is the approach outlined in this paper.

Duller's data sets
Table 1 lists comparisons of the D_e and errors for 3 of Duller's (2007) multi-grain SAR data sets. Duller's (2007) data set D-4 illustrates a linear (L) regression, and the other two, a saturating exponential (E) (samples D-1 and D-2). The E data sets of Duller (2007) are only somewhat sublinear. There is no significant difference in error estimates (within 1σ)

Sample	Model	Equivalent Dose (Gy)		
		Curve Fitting	**Monte Carlo[a]**	**This paper**
D-4	L	0.71 ± 0.11	± 0.13(Sym)	0.73 ± 0.11
D-1	E	0.70 ± 0.06	± 0.06(Sym)	0.698 ± 0.058
D-2	E	28.50 ± 0.67	± 0.75(Sym)	28.48 ± 0.68

[a] 1000 repeats were used. 'Sym' means that the D_e distribution is symmetric. Only the uncertainties are reported because the D_e estimates are those from the curve-fitting scheme

Table 1: *Comparison of equivalent dose values, and 1σ uncertainties, for 3 data sets from Duller (2007).*

Sample	**FUS-1 (1-70)** [a]	**FUS-1 (2-20)**	**FUS-1 (4-40)**		**SFC-6 (A-10)**	**SFC-6 (A-15)**		**ATP-37 (A-1)**
Dose (s)	*L/T*	*L/T*	*L/T*	*Dose (s)*	*L/T*	*L/T*	*Dose (s)*	*L/T*
Natural	1.390 ± 0.378	1.261 ± 0.383	1.703 ± 0.538	Natural	4.957 ± 0.208	2.671 ± 0.212	Natural	4.500 ± 0.107
100	0.091 ± 0.110	0.324 ± 0.139	0.200 ± 0.234	100	1.109 ± 0.053	1.146 ± 0.107	150	0.814 ± 0.021
300	0.352 ± 0.127	0.900 ± 0.455	0.750 ± 0.322	250	2.129 ± 0.092	2.023 ± 0.195	300	1.379 ± 0.034
1000	1.452 ± 0.483	1.909 ± 0.796	1.783 ± 0.649	500	3.278 ± 0.137	2.165 ± 0.178	700	2.360 ± 0.057
0	-0.036 ± 0.037	-0.074 ± 0.057	0.001 ± 0.001	1000	4.306 ± 0.172	3.420 ± 0.288	1200	3.209 ± 0.077
Recycle	1.0 ± 1.6	1.01 ± 0.58	2.4 ± 3.0	2000	5.592 ± 0.226	3.390 ± 0.269	2000	3.925 ± 0.094
				0	-0.023 ± 0.012	-0.022 ± 0.024	2800	4.558 ± 0.108
				Recycle	0.87 ± 0.06	1.07 ± 0.14	3600	4.853 ± 0.114
							4200	5.091 ± 0.120
							5000	5.441 ± 0.128
							5800	5.737 ± 0.135
							7000	6.102 ± 0.143
							0	0.005 ± 0.002
							Recycle	0.88 ± 0.03

[a] In parenthesis, either the 'single'-grain hole (e.g., disc 1, hole 70) or the multi-grain aliquot number. The recycle ratio is for the first applied dose. The FUS and SFC samples are quartz sand grains (respectively, 125-185 μm and 105-212 μm). The ATP-37 sample is 4-11 μm quartz.

Note: Most L/T errors in this table are listed with a precision of 3 significant digits even though it is well known that most data error estimates themselves have uncertainties typically of 10-30% (e.g., Topping, 1962, p. 89-90). The retention of 3 significant digits is intended to minimize any effects of round-off error for those users who wish to employ these data in comparisons of results from their own models with the models in this paper. These L/T ratios are from the screen display of Analyst, which truncates errors to the third decimal place. Note also that the absolute errors in the ATP-37 L/T ratios yield approximately constant relative errors (~2.5%).

Table 2: *Example data for regression models*

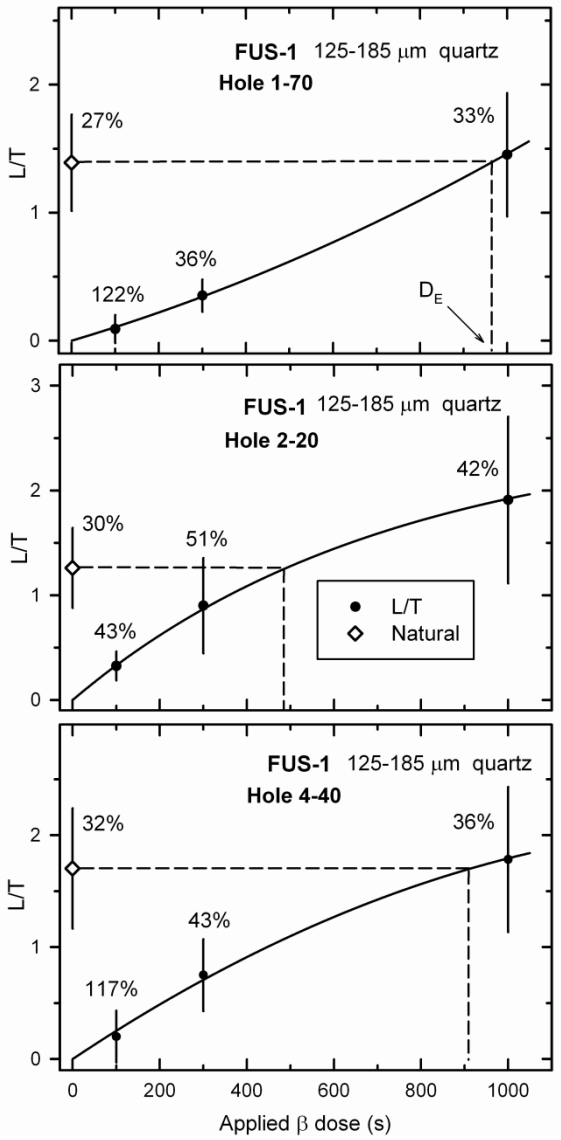

Figure 1: *Top to bottom, Q, E and Q regressions using the routines of this paper for the 'single-grain' FUS-1 data in Table 2. The percent relative error in each L/T ratio is shown.*

among the 3 approaches for the samples D-4, D-1, and D-2. For these data, the fitting parameters (not shown) generated by *Analyst* 2007 and by the author's routines do not differ by more than a small fraction of the 1σ uncertainties in these parameters. That is, the respective best-fit DRCs do not differ noticeably.

Author's data sets

The 6 data sets of the author listed in Table 2 illustrate the use of the Q, E and E+L models. Unlike in Duller's (2007) examples, the dose points in Table 2 are in seconds, not Gy. Other distinctions are in the footnotes to Table 2. The data for the first 3 examples

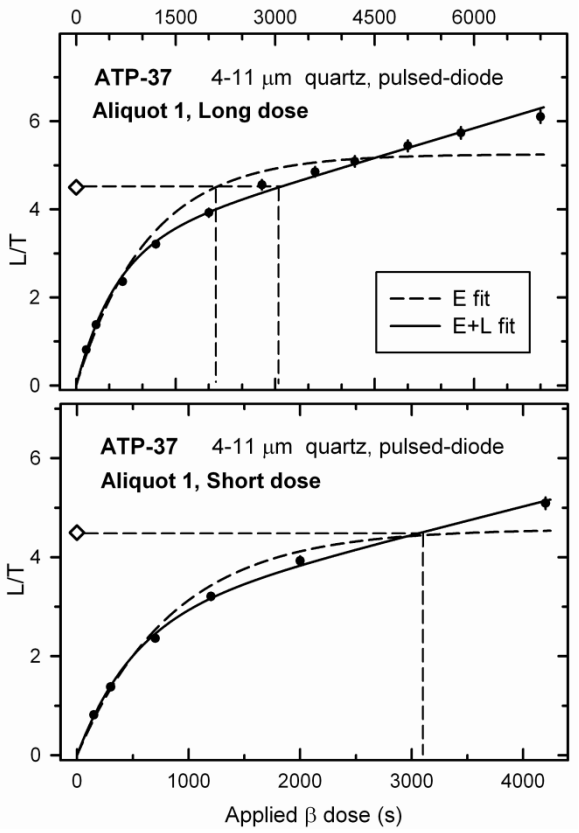

Figure 2: *Comparison of E and E+L best-fit DRCs for sample ATP-37.*

represent SAR conditions for which the relative errors in the L/T ratios are much larger than those in the data sets of Duller (2007). These larger errors manifest the effects of low signals, and so it is likely that the role of systematic errors is larger than would otherwise be the case. Furthermore, these 3 data sets were selected because the scatter of the data about the best-fit DRCs is very small while the absolute and relative errors in the L/T ratios are very large. Finally, the corresponding fitting parameters have very large relative errors (as estimated by *Analyst* 2007).

The data (excluding recycle and recuperation points) for the 3 different grains of the FUS-1 sample are plotted in Fig. 1, with best-fit regressions of the Q (top), E (middle), and Q (bottom) models. The best-fit DRCs of *Analyst* 2007 and of the author do not differ on the scale of these plots. All data have large relative errors. The relative errors in the L_0/T_0 ratios are comparable in all 3 plots. The relative errors in the dose response curve data in the top and bottom plots are comparable, but differ from those in the middle plot. Examples of single-grain data having similarly large relative errors but with more variability within the dose response curve could

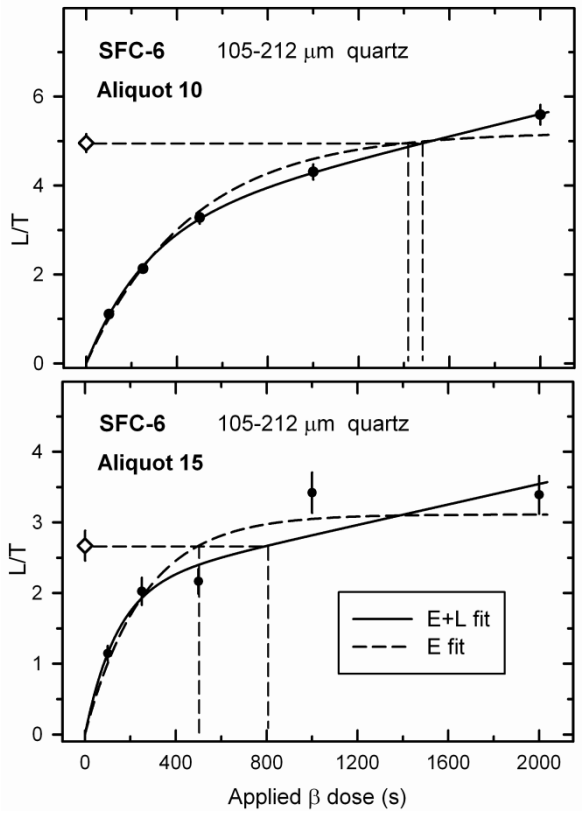

Figure 3: *Comparison of E and E+L best-fit DRCs for sample SFC-6.*

probably be found by the author and by readers, but these examples are sufficient to illustrate some differences in the results from the different error-analyses schemes. The D_e and error results from FUS-1 are discussed below after presenting plots for the remaining data in Table 2.

The multi-grain SAR data for the fine-silt (4-11 μm) quartz of sample ATP-37 are presented not only because they provide DRCs having a nearly 'linear' portion beyond a saturating exponential, but also to illustrate the effects on estimated errors in D_e values of variations in the spacing of dose points. The dose points for the ATP-37 data are approximately equally spaced beyond 1000 s. For extrapolated D_e estimation, the simulations of Grün and Rhodes (1992) showed that doubling of the dose-point spacing (as usually practiced by the author and by BLK87 and B90 for non-SAR data) was preferred to equal spacing of dose points because dose doubling led to smaller errors in the D_e values. However, it is not clear what the effects of different dose-point spacings might be on D_e values and errors calculated for SAR dose response curves (which employ interpolation for D_e estimation). From these ATP-37 data, the dose-doubling scenario can be approximated by eliminating certain dose points in the regressions.

The ATP-37 data for the complete dose range are plotted in the top of Fig. 2, with both E and E+L regressions illustrated. As for the FUS-1 data, the best-fit response curves of the author and of *Analyst* 2007 are indistinguishable at the plotting scale (the fitting parameters agree within their first 3 digits). The bottom of Fig. 2 shows regressions for an approximate 'dose-doubling' scenario, over a truncated dose range (to 4.2 ks, rather than 7 ks).

It is immediately clear in the top of Fig. 2 that the choice of regression model has a dramatic effect on the value of the estimated D_e and that the weighted E model is inappropriate. An equal-weighting (unweighted) E model has not been applied to these data, but it is not expected to provide an acceptable fit. In the lower Fig. 2, a D_e value could not be computed by either *Analyst* 2007's or the author's weighted E-fit routines, because the E fit saturates near the L_0/T_0 value. So in this example incomplete knowledge of the DRC could lead to rejection of the aliquot if only an E model were applied.

The final example data set in Table 2 (sample SFC-6) is a dose-doubling set from a multi-grain dating experiment of the author. The best-fit dose response curves are plotted in Fig. 3. Aliquot 10 (top) represents the form of the dose response for most of the 21 aliquots of this sample: apparently an E+L fit is most appropriate. *Analyst* 2007 provides a 'fit' (estimate of weighted sums of squares of residuals about the best-fit dose response) value of 0.0132 for the E+L model and 0.0959 (~ 7 times larger) for the E model. However, as is shown below, the 3 error-estimation schemes report no statistical difference (at 1σ) in the D_e values from these weighted E and E+L fits.

Aliquot 15 (bottom, Fig. 3) was selected because not only do the data scatter more widely about the best-fit weighted regression dose response, but also the errors in the L/T ratios are much larger than for aliquot 10. In this case, to a first approximation, it could be acceptable to prefer the E+L model based on the assumption that all quartz in this sample should behave the same way. Generally, however, experience shows that quartz dose response curves can vary from grain to grain within a given sample. Therefore, if there are few grains per aliquot, one could expect inter-aliquot differences in dose response.

Sample [a]	Model [b]	Equivalent Dose (s)		
		Curve Fitting	Monte Carlo [c]	This paper
FUS-1/1-70	Q	964±202	±429(Sym)	963±204
FUS-1/2-20	E	491±219	±295(+Asym)	491±213
FUS-1/4-40	Q	914±188	±447(-Asym)	914±493
ATP-37/A-1	E(long)	2104±219	±170(Sym)	2104±128
	E+L(long)	3057±227	±262(Sym)	3056±262
	E(short)	sat'n	-	sat'n
	E+L(short)	3095±185	±224(Sym)	3092±251
SFC-6/A-10	E	1417±769	±171(Sym) [d]	1417±148
	E+L	1489±162	±213(Sym)	1487±172
SFC-6/A-15	E	502±160	±99(Sym)	-- [e]
	E+L	807±299	±314(+Aysm)	806±441

Notes:
a: 'n-m' indicates 'single-grain' disc-hole number; 'A-n' indicates aliquot number.
b: For ATP-37, 'long' means 0-7 ks dose range, while 'short' means 0-4.2 ks.
c: 'Sym' indicates a symmetrical distribution in D_e values, whereas '+Asym' denotes asymmetry (high-side skewness) and '-Asym' denotes low-side skewness. For multi-grain aliquots, 700 repeats are used, whereas 1000 repeats are used for single-hole data. *Analyst* 2007 provides a ±1σ error from the Monte Carlo D_e distribution probably by fitting a Gaussian that is centered on the peak of the distribution's histogram. For an asymmetric distribution, this ±1σ error can be misleading. Only error estimates are reported (see footnote 'a', Table 1)
d: Although symmetrical, the corresponding Monte Carlo distribution peak occurs at ~1200±~150(1σ) s, not 1417, and moreover, only 58% of the repeats are fitted (~55% when 1000 repeats are selected).
e: Does not calculate.

Table 3: *Comparison of $D_e \pm 1\sigma$ results for the data in Table 2.*

The *Analyst* 2007 'fit' values (0.118 for the E model, 0.112 for the E+L model) are probably effectively the same for aliquot 15. Thus even though it might appear in lower Fig. 3 that the E and E+L models provide distinctly different D_e values, the results in Table 3 below indicate that they do not differ at 1σ. Thus aliquot 15 is a good example of an ambiguous dose response curve that could provide misleading D_e values for a sediment sample containing mixed-age quartz grains. In light of the lessons from the plots in Fig. 2, SFC-6 provides an example of a sample for which a greater number of dose points should have been chosen, not necessarily evenly spaced.

Comparison of D_e and errors for the Table 2 data sets

The D_e and $\sigma^2_{D_E}$ values obtained from the regressions shown in Figs. 1, 2 and 3 are listed in Table 3, along with corresponding values from Duller's (2007) 'curve fitting' and Monte Carlo procedures.

The Q-fit error estimates for FUS-1(1-70) from Duller's 'curve-fitting' approach and the author's approach are indistinguishable. Although the relative errors in fitting parameters b and c estimated by *Analyst* 2007 are respectively ~50% and ~200%, this agreement between the two approaches implies that the contributions to the total error in D_e from errors in fitting parameters are negligible in this case. However, both D_e errors under-estimate the Monte Carlo error estimate by a factor of about two. Thus for such data (large relative errors, few dose points, little scatter about the best-fit DRC), one should always check the Monte Carlo result before accepting the error estimates from either of the other two approaches.

With the E-fit DRC of FUS-1(2-20), the two smaller error estimates in D_e do not differ significantly and either of the estimates would imply that this 'grain' should be accepted (error estimates in D_e are <50%). The estimated relative errors in fitting parameters a

and b are ~100%. Thus the agreement of the D_e error (±219 s) from the 'curve-fitting' scheme with that (±213 s) from the author's scheme implies that the effect of errors in the fitting parameters on the total D_e error is negligible in this case also. Nevertheless, a visual inspection of the Monte Carlo D_e distribution suggests that a more accurate error estimate would be 491^{+400}_{-300} s (1σ), and therefore that this 'grain' should be rejected. Caveat emptor!

The error estimates for the Q-fit FUS-1(4-40) also are instructive as to relative performance of the different error-estimation schemes. While the *Analyst* 2007 Q-fit estimate of ±188 s would imply that this D_e value should be accepted, the author's scheme produces a much larger estimate (±493 s), suggesting that a closer look at the data is required. Here the estimated relative errors in the fitting parameters b and c are respectively ~50% and ~200%, similar to those for example FUS-1(1-70). Notwithstanding, the author's error estimate is much closer to the Monte Carlo symmetric estimate (±447 s) from Analyst2007, but even this fact is misleading because the Monte Carlo D_e histogram is highly asymmetric and dramatically skewed to smaller values (negative skewness). A visual inspection of this histogram suggests a more appropriate result of 914^{+150}_{-350} s (1σ). Thus in this example, the Duller 'curve-fitting' error-estimation scheme under-performs, and the other two symmetric-error results (Monte Carlo and author's) schemes over-estimate the asymmetric, most-likely error estimates (+150, -350 s).

The ATP-37 data set also offers some useful insights into relative performance of the 3 error-estimation schemes. The full-range ("long") dose E fits give identical D_e values (2104 s) but apparently different error estimates. From the graph (top Fig. 2) it is apparent that the weighted E fits are not appropriate, and from Table 3 it is clear that the D_e values from the E+L model are significantly larger than the E-fit values (by ~50%). What is unexpected is the relatively small E-fit error estimate (±128 s) produced by the author's scheme. The Monte Carlo error estimate is near the middle of the range of errors produced by the schemes of the author and of *Analyst* 2007. These inter-scheme differences for the E model may reflect effects from the use of weighting, but that is a subtlety requiring more study (as suggested by the reviewer, perhaps using numerical simulations, e.g., Grün and Rhodes, 1992). Notwithstanding, the E+L model yields essentially identical results from the 3 approaches, with Analyst2007 giving a 15% (statistically perhaps insignificant) under-estimation of the error (compared to the Monte Carlo result). Is this apparent under-estimation a consequence of not capturing the contributions from the errors and covariances in the fitting parameters? Perhaps only model simulations could answer this question.

Concerning the short-dose regressions, because the L_0/T_0 ratio is at or near the saturation value of the exponential fit, neither *Analyst* 2007 nor the author's scheme could calculate a D_e value from the E model. On the other hand, both schemes produced effectively identical D_e values from the use of the E+L model. Where they differ in this case is that again *Analyst* 2007's 'curve-fitting' scheme yields the lowest error estimate. Again, this may reflect the neglect of the contributions from the errors and covariances in the fitting parameters. What is also helpful to notice is that any effect of changes in the dose-point spacing (compare top and bottom of Fig. 2) on E+L results when the L_0/T_0 intersection is well within the 'linear' part of the dose response is not statistically resolvable with this data set. This suggests that with SAR, the type of dose-point spacing may not be a significant variable in the estimation of D_e and its error, as long as the scatter about the best-fit E+L DRC is small.

The final data set (SFC-6) provides an example of dose responses for which a critical variable may be the number of dose points, not the spacing of same. Aliquot 10 of sample SFC-6 apparently yields essentially identical D_e values whether the E or E+L models are used. However, the error estimates from the use of the E model differ among the 3 schemes, and the *Analyst* 2007 E-fit Monte Carlo D_e histogram is anomalous (footnote 'd'). This Monte Carlo routine did not execute to completion (footnote 'd') so the asymmetric D_e histogram (peak at ~ 1200 s) provides no helpful information about what the statistically realistic D_e error should be. One could certainly expect an asymmetric Monte Carlo D_e distribution from this E-fit example, but a symmetric result is generated. The *Analyst* 2007 'curve-fitting' error estimate of ±769 s for this E fit would appear (from Fig. 3) to be a more reasonable error estimate than those from the other two approaches. For aliquot 10, only the weighted E+L model yields general consistency among the 3 error-analysis schemes, and thus this model is preferred for this aliquot.

As mentioned above, aliquot 15 was selected to represent a dose response having both a larger scatter of L/T ratios about the DRC and larger relative errors (about double) in the L/T ratios than those for aliquot 10. Given the nearness of the L_0/T_0 ratio to the E-fit saturation and the paucity of the dose points, it is surprising how small the error estimates are from the *Analyst* 2007 'curve-fitting' and Monte Carlo

schemes. Neither seems realistic. On the other hand, all 3 schemes yield E+L model error estimates that overlap at 1σ with the E-fit D_e estimates.

Notwithstanding, there are differences among the E+L error estimates, with the *Analyst* 2007 'curve-fitting' providing the smallest estimate and the author's scheme, the largest. This largest estimate (±441) appears to reflect more accurately the large spread (at 2σ) of D_e values within the Monte Carlo histogram than does the 'curve-fitting' error estimate of *Analyst* 2007. However, only the Monte Carlo histogram for the E+L results from aliquot 15 appears to produce a realistic estimate in this case. In particular, this histogram suggests a D_e value of 807^{+300}_{-200} s (1σ). Aliquot 15 yields somewhat ambiguous model results, and without obtaining additional dose points, it is acceptable (to a first approximation) to prefer the E+L model because that seems to be applicable to most of the other aliquots (not shown) from this sample. It is clear that the author's error-analysis scheme could lead to the rejection (D_e error >50%) of this aliquot, whereas the other schemes would not.

Conclusions

Duller's (2007) 'curve-fitting' approach to estimation of errors in SAR D_e values, executed in *Analyst* 2007 software, does not capture the contributions from errors in the regression parameters nor from any covariances among the errors in these parameters. The author's error-analysis schemes do capture these contributions. However, comparisons of the computed D_e values and their error estimates from these two schemes applied to selected data sets generally show no statistically significant differences in error estimates except in special cases.

Moreover, the Monte Carlo scheme executed in *Analyst* 2007 can in some cases indicate a D_e histogram peak at a value significantly different from the 'central D_e value' reported by *Analyst* 2007.

Generally, however, for SAR data that have relatively small scatter about the best-fit dose-response curves and that have relative errors in L/T ratios smaller than ~5%, there are no significant differences among the 3 discussed schemes for analysis of L, Q, E and E+L models.

Nevertheless, some data sets (some presented here) can generate misleading D_e and error estimates, and the practitioner of SAR dating should inspect such data sets carefully before accepting D_e and error estimates, no matter which error-analysis scheme is employed.

Finally, one selected data set (ATP-37, Table 2) provides a good example of quartz SAR E+L dose response, and is used here to demonstrate that there is (in this case) little or no dependence of computed D_e values and their error estimates on dose-point spacing schemes, whether spaced evenly or by dose-doubling. This contrasts with the dose-point-spacing dependency shown by the simulations of Grün and Rhodes (1992) for non-SAR data (for which extrapolation is used).

Acknowledgements

Constructive review comments on the first two drafts of this manuscript by Dr. Kristina Thomsen are greatly appreciated.

References

Ballarini, M., Wallinga, J., Wintle, A.G., Bos, A.J.J. (2007) Analysis of equivalent-dose distributions for single grains of quartz from modern deposits. *Quaternary Geochronology* **2**, 77-82.

Berger, G.W. (1987) Thermoluminescence dating of the Pleistocene Old Crow tephra and adjacent loess, near Fairbanks, Alaska. *Canadian Journal of Earth Sciences* **24**, 1975-1984.

Berger, G.W. (1990) Regression and error analysis for a saturating-exponential-plus-linear model. *Ancient TL* **8**, 23-25.

Berger, G.W. (2009) Zeroing Tests of Luminescence sediment dating in the Arctic Ocean: Review and new results from Alaska-margin core tops and central-ocean dirty sea ice. *Global and Planetary Change* **68**, 48-57.

Berger, G.W., Lockhart, R.A., Kuo, J. (1987) Regression and error analysis applied to the dose response curves in thermoluminescence dating. *Nuclear Tracks and Radiation Measurements* **13**, 177-184.

Berger, G.W., Sawyer, T.L., Unruh, J.R. (2010) Single- and multigrain luminescence dating of sediments related to the Greenville fault, eastern San Francisco Bay area, California. *Bulletin of the Seismological Society of America* **100**, 1051-1072.

Bluszcz, A. (1988) The Monte Carlo experiment with the least squares methods of line fitting. *Nuclear tracks and Radiation Measurements* **14**, 355-360.

Cunningham, A. C., Wallinga, J. (2010) Selection of integration time intervals for quartz OSL decay curves. *Quaternary Geochronology* **5**, 657-666.

Deming, W.E. (1964) *Statistical Adjustment of Data*. Dover Publications, Inc., New York, 255pp.

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

Galbraith, R. F. (2002) A note on the variance of a background-corrected OSL count. *Ancient TL* **20**, 49-51.

Grün, R., Rhodes, E.J. (1992) Simulation of saturating exponential ESR/TL dose response curves -- weighting of intensity values by inverse variance. *Ancient TL* **10**, 50-56.

Grün, R., Packman, S. (1993) Uncertainties involved in the measurement of TL intensities. *Ancient TL* **11**, 14-20.

Hayes, R. B., Haskell, E. H., Kenner, G. H. (1998) An assessment of the Levenberg-Marquardt fitting algorithm on saturating exponential data sets. *Ancient TL* **16**, 57-62.

Lai, Z. P. (2010) Chronology and the upper dating limit for loess samples from Luochuan section in the Chinese Loess Plateau using quartz SAR protocol. *Journal of Asian Earth Sciences* **37**, 176-185.

McCalla, T.R. (1967) *Introduction to Numerical methods and FORTRAN Programming*. John Wiley & Sons, New York, 351pp.

Murray, A. S., Wintle, A. G. (2003) Luminescence dating of quartz using an improved single-aliquot regenerative-dose protocol. *Radiation Measurements* **37**, 377-381.

Murray, A. S., Svendsen, J. I., Mangerud, J., Astakhov, V. I. (2007) Testing the accuracy of quartz OSL dating using a known-age Eemian site on the river Sula, northern Russia. *Quaternary Geochronology* **2**, 102-109.

Murray, A. S., Buylaert, J-P., Henriksen, M., Svendsen, J-I., Mangerud, J. (2008) Testing the reliability of quartz OSL ages beyond the Eemian. *Radiation Measurements* **43**, 776-780.

Pawley, S. M., Bailey, R. M., Rose, J., Moorlock, B. S. P., Hamblin, R. J. O., Booth, S. J., Lee, J. R. (2008) Age limits on Middle Pleistocene glacial sediments from OSL dating, north Norfolk, UK. *Quaternary Science Reviews* **27**, 1363-1377.

Press, W. H., Flannery, B. P., Teukolsky, S. A., Vetterling, W. T. (1986). *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press.

Topping, J. (1962) *Errors of Observation and their Treatment*. Chapman and Hall, London, 116pp.

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.

Reviewer
K. Thomsen