Observations on palæodose determination with burnt flints

N. Mercier, H. Valladas and G. Valladas
Centre des Faibles Radioactivités, Lab. mixte CNRS-CEA, Ave. de la Terrasse, 91198 Gif-sur- Yvette Cedex, France.

Introduction

The thermoluminescence dating of prehistoric flints is now a solidly established technique (Aitken, 1985; Huxtable and Aitken, 1986; Valladas, 1992). However, in attempting to date Lower or Middle paleolithic hearths one is confronted by three problems:

1. The supralinear growth of the TL within a wide interval of radiation doses, up to 100 Gy in some instances (Mercier, 1991);
2. the onset of saturation at doses in the 150 to 700 Gy range;
3. the specific behaviour of the α radiation, which contributes significantly to the integrated radiation dose in the flint.

The first two are associated with the β and γ components of the radiation dose.

This paper deals with procedures for determining palaeodoses in archaeological flints when faced with obstacles of the type noted above and describes simulation experiments on raw unburnt flints collected near palaeolithic hearths. The experimental conditions are listed in table 1.

Table 1. Experimental conditions

The samples were annealed at 500 °C (except when indicated otherwise), crushed in an agate mortar and sieved to produce a size range of 100-160 μm, followed by washing with hydrochloric acid to remove calcium carbonate. TL emission was detected between 350 and 450 nm and glow curve measurements were made with six 2 mg aliquots for each sample.		
The radiation sources were:		
α	^{238}Pu	Flux (4MeV): $2.5 \cdot 10^6$ α/(cm^2·sec)
β	^{90}Sr	Dose rate: ~8 Gy/min
γ	^{137}Cs	Dose rate: ~100 Gy/hr with 50 mg samples (Valladas, 1978).

In the discussion that follows the effects of β and γ radiation will be considered as equivalent.

The effects of α radiation.

It has been reported that the TL induced by α–particles is a linear function of the integrated α–flux (Aitken, 1985). However, it seemed advisable to check whether the linearity of the α–contribution (TLN_α) is maintained in flints heated 300 ka - 400 ka ago whose TLN ($TLN = TLN_\alpha + TLN_\beta$) indicates the onset of saturation. Thus two specimens (labelled T53 and T59) were subjected to the fine grain preparation (Zimmerman, 1971) and irradiated. As can be seen in fig.1 the growth curves remain linear up to integrated α fluxes of ~2.7 and ~$0.9 \cdot 10^{10}$ α/cm^2 for samples T53 and T59 respectively. Based on the uranium content of the two flints, the corresponding onset of saturation is expected after 1.14 and 1.67 million years respectively. Thus we will assume in this paper that the growth curve is a linear function of the integrated α–flux.

The effects of α and β (or γ) irradiation

To determine the relative efficiency of α and β (or γ) rays in the production of TL at different levels of saturation (Aitken, 1984), we worked on single layers of grains from prehistoric burnt flints T8 and T60 treated as described in table 1, except for the initial heating. The short path length of α rays compared with the grain size is unimportant since we are only interested in the relative effects of the two radiations. Four samples of each specimen were first exposed to γ radiation doses of 0, 150, 300 and 600 Gy respectively. A fraction of each sample received an additional $9 \cdot 10^9$ α/cm^2; another fraction received an additional β dose of 126 Gy.

The TL growth induced by the α and β radiation is shown in fig.2 (ΔTLα and ΔTLβ, respectively) as a function of each of the four initial γ doses. It can be seen that the higher the initial γ dose the smaller the subsequent ΔTLβ; such behaviour is indicative of saturation. In contrast, the ΔTLα is almost independent of the degree of saturation. Thus, the α/β efficiency ratio rises with increasing γ (or β) dose. This has to be taken into consideration when computing the β dose equivalent to a given α irradiation, which is:

$$S_\alpha = \frac{(TL_\alpha)D_\beta}{(TL_\beta)\,\varphi_\alpha} \quad Gy/(\alpha/cm^2) \qquad (1)$$

where,
$TL_\alpha = TL(N + \alpha) - TLN$; $TL_\beta = TL(N + \beta) - TLN$,
D_β is the β dose received by the sample and φ_α is the integrated α flux.

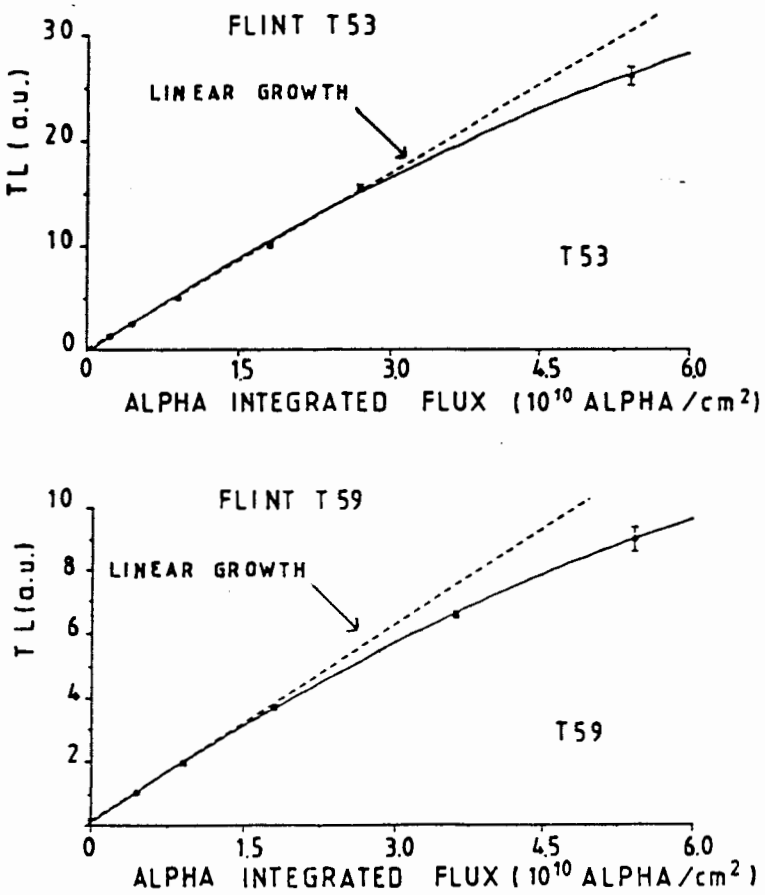


Figure 1.
TL growth in flints T53 and T59 exposed to α radiation only. The solid lines represent polynomial computer fits through the experimental points.

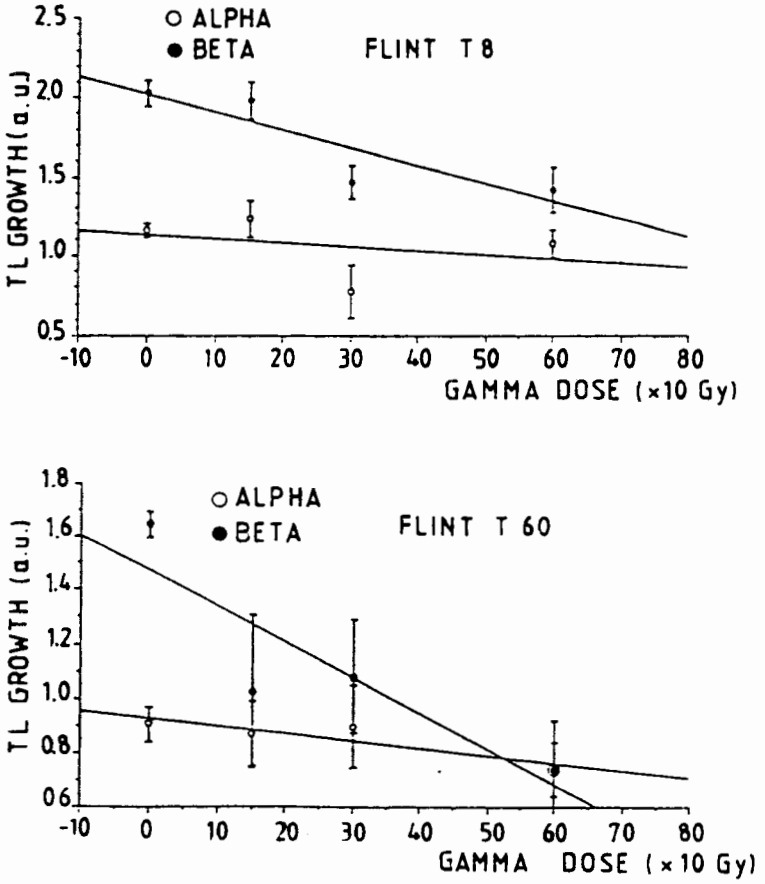


Figure 2.
TL growth in prehistoric burnt flints T8 and T60 receiving, respectively, an integrated flux of 9·10⁹ α/cm² (ΔTLα) and a β dose of 126 Gy (ΔTLβ) as a function of the initial γ dose received (0, 150, 300 and 600 Gy).

It is apparent that for archaeological flints saturation will result in an overestimated S_α. One way of getting around this difficulty is to determine S_α on annealed samples. Tests on several flints which were not saturated showed that the value of S_α is unaffected by annealing at 500 °C. Hence, for samples at the onset of saturation, S_α should be determined in this manner.

Palaeodose determination in flints at the onset of saturation

The induced TL of an archaeological flint is normally plotted as a function of the added dose. Finding a function that passes through the experimental points and can be extrapolated to zero to give the palaeodose can present problems. The function currently used to deal with saturation has the form:

$$A\,(1 - e^{-d/M})$$

where, A is a proportionality factor, d represents the dose and M is a parameter giving the best fit to the experimental curve.

However, measurements on flints heated in the laboratory show that for d>0, the growth of TL is parabolic (i.e. $ad + bd^2$, where b>0) over a relatively wide dose range (0 to 100 Gy) which corresponds to supralinear behaviour (fig.3 in Mercier, 1991). It is noteworthy that expansion of the exponential formula gives $A(d - d^2/2M^2)$, a parabola with a curvature inverse of the experimental findings and clearly inappropriate. When extrapolated, such an exponential usually underestimates the palaeodose by about 50 Gy (fig.3 in Mercier, 1991).

We failed to find a simple function capable of fitting TL growth curves within the dose interval of interest for dating measurements. The only practical way to compute the palaeodose of flints exhibiting supralinearity and saturation is to use the glow curve recorded during second heating. Such an approach does not require any knowledge of the missing parts of the first TL glow curve, as has been shown in a previous study of quartz pebbles heated in a volcanic lava flow (Valladas and Gillot, 1978). However, in the case of flints, the difference between the plots of the TL growth induced by the α and β radiations must be taken into account. The procedure is applicable if the TL glow obtained after the second heating has values proportional to those obtained during the first heating. If the condition is fulfilled the first glow curve can be normalized to the second in the manner described below.

Let TL(D) represent the natural TL and TL'(D') the second TL glow, where D' is the applied laboratory dose under conditions where TL' (D') ≈ TL(D).

Let TL(D + d) and TL' (D' + d) represent the first and second glow after a dose, d, has been added.

We now define:

$$Y_o = \frac{TL(D)}{TL'\,(D')} \cdot (D') \qquad (2)$$

$$Y_1 = \frac{TL(D+d)}{TL'\,(D'+d)} \cdot (D'+d) \qquad (3)$$

$$Y_2 = \frac{TL(D+2d)}{TL'\,(D'+2d)} \cdot (D'+2d) \qquad (4)$$

To get the first estimate of the palaeodose, D_o, one extrapolates the points $(0,Y_o)$, (d,Y_1), .which usually yields a straight line, in the normal way (see fig. 7 in Valladas and Gillot, 1978). The second estimate, D_1, is obtained by replacing D' with D_o in equations (2), (3), etc. After 3 or 4 iterations the series D_o, D_1, D_2, ... converges to the correct palaeodose, D. The ratios

$$\frac{TL(D)}{TL'(D')}\,,\ \frac{TL(D+d)}{TL'(D'+d)}\,,\ \cdots$$

are then independent of glow curve temperature in the temperature range of the TL peak, because the doses D and D', (D+d) and (D'+d) are not significantly different and converge at the end of iteration.

Measurements were made to check if the flints satisfied the proportionality condition. Raw flints collected near prehistoric hearths were heated to 500 °C, within the temperature range estimated for prehistoric burnt flints (Valladas, 1983) and under the conditions given in table 1. Half of the grains prepared taken to represent flint freshly removed from the hearth, was set aside. The other half was annealed at 350 °C for 90 mins after receiving a dose of 300 Gy to simulate the palaeodose. The annealing was performed at a temperature slightly lower than that of the TL peak (380 °C) so as not to alter the TL properties of the second flint fraction in comparison with properties of the first. The TL of the first half was denoted by TL1, that of the second by TL2. Aliquots of both fractions were irradiated with doses of 50, 100, 250, 400, 600 and 800 Gy, respectively, and their TL measured. The corresponding TL1 and TL2 values were obtained by integration of the glow peak between 350 and 410 °C. For the different applied doses the TL1/TL2 ratios were computed from polynomials fitted to the experimental values for eight flints: A, B, C, D, E, H, I and J (see fig. 3). As can be seen, the ratios are relatively constant for flints B, C, D, E, I and J and vary by less than 5% in flints A and H.

Since the proportionality condition is fairly well satisfied, the method described above can be used to obtain a good estimate of the palaeodose in the following way: 1) determine the growth curve of one fraction of the grains of archaeological flint; 2) determine the growth curve of another fraction, annealed at 350 °C, at doses higher than the roughly estimated paleodose. Interpolation of a second order polynomial is generally sufficient. It should be pointed out that the

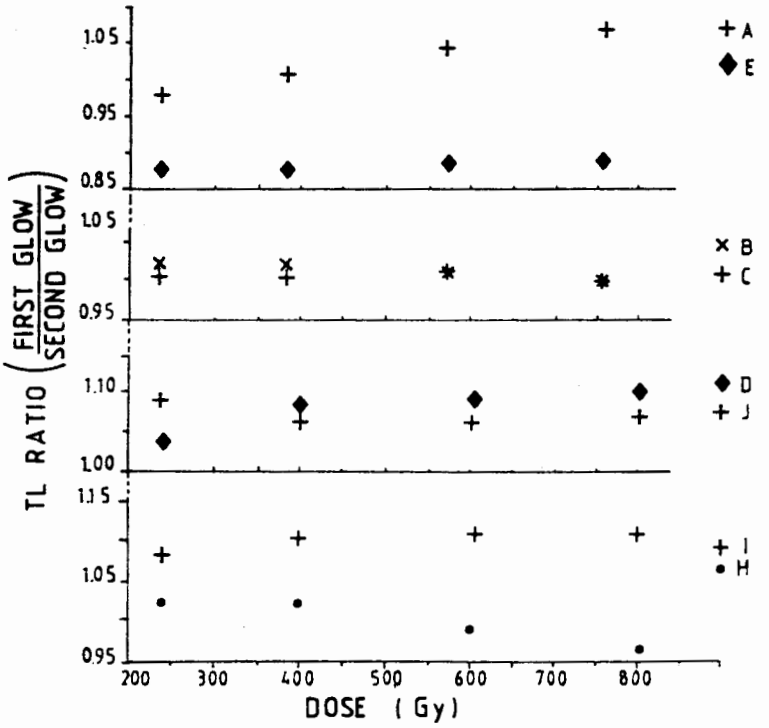
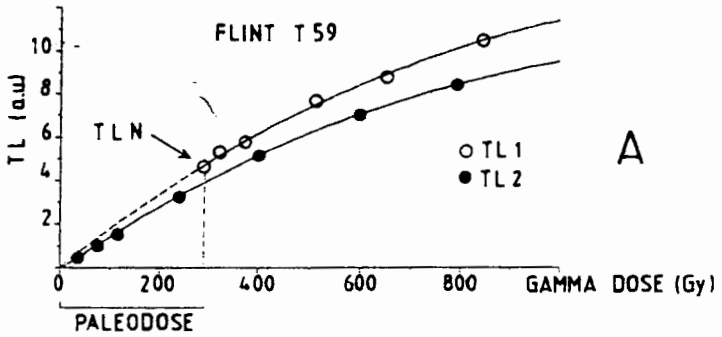


Figure 3.
Ratios of TL values from the first and second glows as a function of irradiation dose in experimental flints A, B, C, D, E, H, I and J.



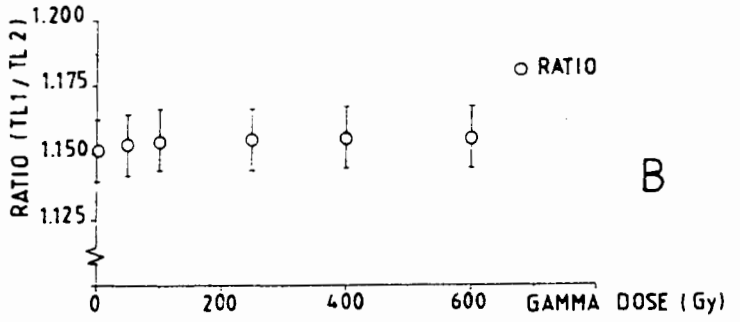


Figure 4.
Palaeodose determination in archaeological flint T59.

A. TL growth for the first (TL1) and second (TL2) heatings, respectively, as a function of the added γ dose. Each point represents a value with an uncertainty of less than 3%. The solid lines represent polynomial computer fits through experimental points.

B. Ratios between corresponding points on TL growth curves from the first and second heatings as a function of γ radiation dose added.

result is independent of the flint response at lower doses.

All the above concerns primarily flints that have received β (or γ) radiation. However, the TLN of prehistoric flints invariably contains some contribution from α radiation. As the proportionality condition has been found to be satisfactory with experimental β (or γ) irradiated flints, it is reasonable to work with prehistoric flints in the same way and subtract the α–contribution from the TLN. The palaeodose induced by β or γ rays from which the "β age" is computed by using the equation:

$$A = \frac{\text{Palaeodose}_{(\beta+\gamma)}}{\text{Annual dose}_{(\beta+\gamma)}}$$

The α–contribution to the TLN can be approximated by:

$$(TLN)\alpha = \frac{(TLN) \cdot d_\alpha}{\text{Annual dose}}$$

where, d_α has the form: $d_\alpha = S_\alpha \cdot$ (annual α flux) and S_α is given by formula (1).

If the TLN is at the onset of saturation the age thus computed will be overestimated. In that case, a new $(TLN)_\alpha$ can be calculated by using this approximated age and reiterating the calculations. In most instances the age thus recalculated falls within 2% of the true age.

A good example of this can be seen in fig.4, where the data from a prehistoric burnt flint are plotted. Part A shows the TL growth as a function of radiation dose during the first and second heating, respectively and the determination of the palaeodose. In part B of fig. 4, the relative constancy of ratios calculated by interpolating the two polynomials used to express the TL growth during the first and second heating respectively is evident. The radiation received by the flint ranged from the palaeodose (point 0 on the abscissa) to 600 Gy of additional laboratory dose.

Conclusion

We have shown that a hypothetical TL growth curve is not necessary to determine the palaeodose of burnt flints at the onset of saturation. After the α-contribution to the natural TL has been subtracted a β-palaeodose can be computed from glow curves recorded after second heating.

References

Aitken, M. J. (1984) Non-linear growth: allowance for alpha particle contribution. *Ancient TL* **2**, 2-5.

Aitken, M. J. (1985) *Thermoluminescence Dating*. Academic Press, London, 191-196.

Huxtable, J. and Aitken, M. J. (1986) Dating of European flint by thermoluminescence, Proc. 24th Archaeometry Symposium (J. S. Olin and M. J. Blackman, eds.), Smithsonian Institution, Washington DC, 465-472.

Mercier, N. (1991) Flint paleodose determination at the onset of saturation. *Nucl. Tracks Radiat. Meas.* **18**, 77-79 .

Valladas, G. (1979) A gamma ray irradiator. *PACT* **3**, 439-442 .

Valladas, G. and Gillot,Y. (1978) Dating of the Olby lava flow using heated quartz pebbles: some problems, *PACT* **2**, 141-150.

Valladas, H. (1978) Thermoluminescence dating of burnt stones from a prehistoric site. *PACT* **2**, 180-183.

Valladas, H. (1983) Estimation de la température de chauffe des silex préhistoriques par leur thermoluminescence. *C. R. Acad. Sci. Paris* **296**, II, 993-996.

Valladas, H. (1992) Thermoluminescence dating of flint. *Quaternary Science Reviews* **11**, 1-15.

Valladas, H. and Valladas, G. (1982) Effet de l'irradiation alpha sur les grains de quartz, *PACT* **6**, 171-178.

Zimmermann, D. W. (1971) Thermoluminescent dating using fine grains from pottery. *Archaeometry* **13**, 29-52.

PR Vagn Mejdahl

TL dating in the Holocene using red TL from quartz

M. Montret, D. Miallier, S. Sanzelle, J. Faïn, Th. Pilleyre and S. Soumana.
Laboratoire de Physique Corpusculaire, IN2P3 - CNRS et Université Blaise Pascal, F 63177, Aubière Cedex, France.

Introduction

With a view to dating Quaternary volcanoes, our group has studied the red TL of quartz observed by Hashimoto and Habuki (1987) with quartz grains extracted from Japanese volcanic materials and by Huntley et al. (1988) in quartz grains from southern Australian beaches. More information was obtained for a peculiar orange-red peak, around 360 - 390 °C (5 °C/s heating rate) and 610 - 620 nm, which was observed in almost all the samples, from various origins, tested in our laboratory (Miallier et al., 1991). Several red TL ages of volcanic events between 14 and 800 ka were subsequently measured (Pilleyre, 1991; Pilleyre et al., 1992). At the same time, Hashimoto et al. (1991) applied the red TL to dating tephra layers correlated with the oldest prehistoric sites in Japan and obtained four ages in the range 60-200 ka.

In the work reported in this paper, we tested the red TL method on two known age samples younger than 10 ka, a brick from a Celtic potter's kiln excavated at Bas-et-Lezat (Puy de Dôme, France) and archaeologically dated to 120 ± 40 BC and a soil baked by a lava flow from the La Vache - Lassolas twin volcanoes (Chaîne des Puys, French Massif-Central) around 8.6 ka ago (sampled beneath the Saint Saturnin Castle). A ^{14}C date of 7,650 ± 350 BP (Sa-90) was produced for carbonized wood sampled below the lava-flow (Pelletier et al., 1959) and another ^{14}C date of 7,970 ± 125 BP (MBN-328) was obtained on a peat containing a tephra attributed to the same event by Juvigné and Gewelt (1987). These results fall just outside the limits of the currently available calibration curve; however, all data indicate that the difference between the actual age and the ^{14}C age is around 800-900 years for this period (Stuiver et al., 1986; Vogel, 1987; Bard et al.,1990). Huxtable et al. (1978) obtained a mean TL age of 8,100 ± 800 a using the fine grain and quartz inclusion techniques on the sediment baked by the lava-flow; Guérin (1983) obtained two TL ages for plagioclase grains extracted from the lava itself, of 9,150 ± 550 a and 8,820 ± 870 a, respectively. The weighted mean for these ages is 8,650 ± 450 a (1σ), although it is not possible to average rigorously because of the specific constraints of the ^{14}C technique.

Technique

The gamma annual dose rate was measured in situ using a portable gamma spectrometer (Sanzelle et al., 1988) and the cosmic radiation dose was calculated on the basis of the localization and depth of the samples (Prescott and Stephan, 1982); the beta dose rate from ^{40}K was calculated from potassium contents and the beta dose rate from the U and Th series was derived from thick source alpha counting (TSAC). The internal beta-equivalent dose rate within quartz grains was assumed to be 0.06 ± 0.05 Gy/ka (from the mean contents of U and Th of quartz grains from the same area , and assuming a value of 0.1 for the efficiency of alpha particles as measured using the red TL from quartz. Dose rate values are given in Table 1.

Quartz inclusions were prepared using a procedure based on Fleming's (1970) recommendations. It included gentle crushing, sieving, magnetic separation, hydrofluoric and hexafluorosilicic acid etching and heavy liquid separation. As a precaution, the preparation was carried out under subdued light, although the red TL of quartz seems to be insensitive to sunlight bleaching .

The irradiations were performed with a ^{137}Cs gamma source delivering 38 mGy/s to quartz. TL measurements were made at a heating rate of 5 °C/s in a nitrogen atmosphere, with - or without - preliminary evacuation. We have found that the red TL is liable to be accompanied by a significant spurious signal if the grains have been crushed or if a long time (several months) elapsed between preparation and measurements. However, this spurious TL can be strongly diminished by gentle acid etching and evacuation.

A longpass sharp cut-off filter Schott RG 610 (50% at 610 nm) was used with a bi-alkali EMI 9635QA photomultiplier tube. The use of a red extended PM tube appeared unnecessary; we observed no gain with such a tube (EMI 9558, kindly lent by the Oxford Laboratory), the lower limit of the technique being controlled by the signal-to-noise ratio.

The additive dose technique was used, with quartz grain aliquots of around 9 mg calibrated by volume. For the two samples, a 'second reading' growth curve was obtained with a laboratory reset (380 °C for 16h in air) natural sample .

Results

The brick. Although we had previously thought that around 10 ka would be the lower limit for the red TL to be measurable (Miallier et al., 1991), a red TL peak was obtained for samples from the brick (age about 2,100 a) after a careful background subtraction (figure 2). The signal-to-noise ratio for the natural signal was ~1 at the peak temperature.

This sample provides an example of the small scatter of the red TL peak: the sample to sample variation of the natural red peak was 6%, whereas it was 27% for the blue signal (figure 1), measured through a blue band-pass Leitz BG 12 filter (the weight dispersion, using the calibrated volume with 200-315 µm grains is usually 2% and this technique results typically in a standard deviation of 1-2% for the TL of good phosphors such as Al_2O_3 or $CaSO_4$:Dy). No difference was found between the curves recorded with or without evacuation. Thus, the red TL appears better distributed than the blue TL and probably less affected by feldspar remains and surface phenomena - which might explain the beneficial effect of HF etching observed by Franklin and Hornyak (1992); moreover, like these authors, we had observed, with samples from another origin, that most of the blue TL could be emitted by less than 10% of the grains (Ousmoï, 1989) and our efforts to normalize with the pre-dose peak were unsuccessful. In the case of the brick, we did not try the dose normalization with the blue filter; this technique can slightly reduce the scatter of the results when the changes of sensitivity are taken

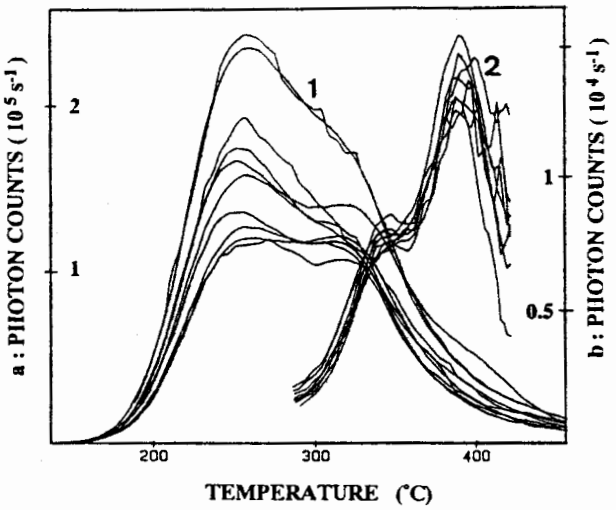


Figure 1. (above)
Individual glow-curves (background substracted) for the quartz of the brick measured using: a) blue filter; b) red filter under conditions described in text.

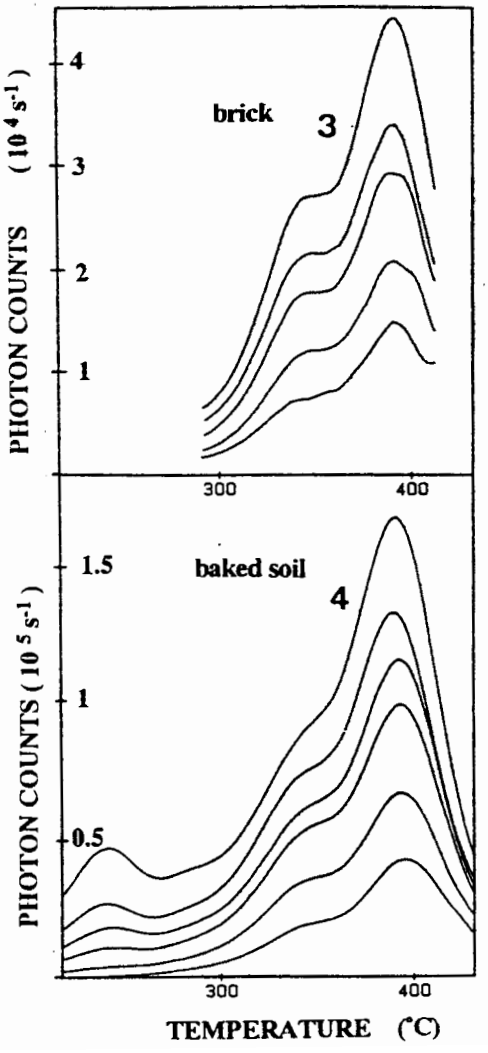


Figure 2. (opposite)
Glow curves for: a) brick and b) baked soil. The background was subtracted and the curves were averaged over 8 or 10 curves for each dose. The curves shown are, from the lowest intensity curve (natural), those obtained after administration of the following gamma doses: a) 4.53; 9.06; 13.6; 18.13 Gy and b) 22.66; 45.32; 67.98; 90.64, and 136 Gy.

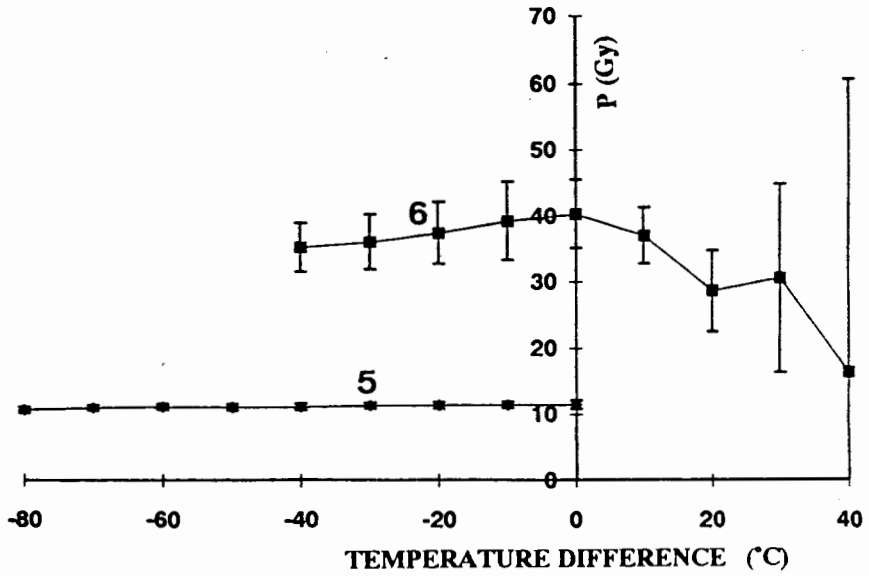


Figure 3.
Palaeodose plateau tests: a) brick; b) baked soil. Above the peak temperature (0 on the X axis), results are less reliable because of the predominance of the thermal background.

Table 1. Summary of results. Quoted errors correspond to 1σ. The evaluations of E and Q were based on the measurement of the peak intensity.

Sample	Grain size µm	E Gy	Q Gy	P Gy	Annual dose Gy/ka	Red TL Age ka	Expected age ka
Brick kiln (Cler 71)	200-315	2.41±0.36	9.11±0.53	11.52±0.61	4.69±0.28	2.46±0.20	2.11±0.04
Baked soil (Cler 207a)	100-200	2.83±0.48	37.5±5.2	40.3±2.3	4.32±0.19	9.34±0.67	8.65±0.45
Baked soil (Cler 207b)	200-315	3.65±0.61	33.7±2.4	37.4±2.5	4.19±0.19	8.92±0.72	8.65±0.45

into account (see Reviewers Comments by J.R. Prescott on the paper by Franklin and Hornyak, 1992).

The growth of the signals vs added dose was linear and the peak did not shift significantly towards low temperatures with increasing dose as is the case at higher doses (fig. 2). A sensitization of 1.3 was found following the second heating. The intercept correction E (terminology from Aitken, 1992) was not negligible and varied with temperature, as did the equivalent dose Q. A good palaeodose plateau (P=E+Q) was obtained (fig. 3), leading to the age (see table 1):

A (Cler71) = 2.46 ± 0.20 ka

The baked soil. The behaviour of the red TL was similar to that of the first sample, but for the shape of the growth curve, which was non-linear (saturating exponential) at high doses. A small shift of the peak with additional dose was indicated but was neglected within error limits (±3 °C). The sensitization was negligible and no change in shape was observed between the first and the second growth curves; an acceptable palaeodose plateau was obtained (fig. 3). The age, representing the average of the results of two grain sizes was (see table 1):

A (Cler 207 a/b) = 9.13 ± 0.72 ka

Conclusion

The two tests reported in this paper show that the red TL peak of quartz can be expected to give acceptable age results between 2 and 9 ka (table 1). This preliminary evaluation should not be extrapolated to older ages because new features will appear at higher doses such as (i) a significant shift of the peak with accrued dose and (ii) generalization of sublinear dose growth-curves which cause difficulties in extrapolations and in the estimation of error. These two topics will be discussed in forthcoming papers. Also, at high doses, low temperature peaks can more.or less be superimposed on the 'good one' (~380 °C) and preheating can be used to

isolate the peak. The conclusion of the present work that might be more securely extended to older samples is the confirmation (of previous work on the red TL of quartz) that anomalous fading and dose rate effects are either weak or negligible.

For Holocene samples the technique described in this paper is similar to the conventional quartz inclusion technique, the main drawback being the increasing predominance of the thermal background. However, to those TL workers faced with the common problems encountered with the quartz inclusion technique, such as poor reproducibility of the glow-curves and presence of sensitization in the low-temperature peaks overlapping those used for dosimetry, we recommend exploration of the use of red TL to reduce these problems.

References

Aitken, M.J. (1992) Symbols in TL and optical dating: provisional list. *Ancient TL* **10**, 15-16.

Bard, E., Hamelin, B., Fairbanks, R. G. and Zindler, A. (1990) Calibration of the ^{14}C timescale over the past 30,000 years using mass spectrometric U-Th ages from Barbados corals. *Nature* **345**, 405-410.

Fleming, S.J. (1970) Thermoluminescence dating. Refinement of the quartz inclusion method. *Archaeometry* **12**, 133-147.

Franklin, A.D. and Hornyak, W.F. (1992) Normalization of inclusion size quartz TL data. *Ancient TL* **10**, 1-6.

Guérin, G. (1983) Thermoluminescence des plagioclases. Méthode de datation du volcanisme. Thesis, Paris, Univ. Pierre et Marie Curie, 258p.

Hashimoto, T. and Habuki H. (1987) Emission properties of thermoluminescence from natural quartz. Blue and red response to absorbed dose. *Nucl. Tracks Radiat. Measts* **13(1)**, 216-218.

Hashimoto, T., Kojima, M. and Sakai, T. (1991) Age determination of prehistorical sites in Japan using red-thermoluminescence measurements from quartz grains. Archaeology and Natural Science, **23**, 13-25 (in Japanese, abstract in english).

Huntley, D.J., Godfrey-Smith, D.I., Thewalt, M.L.W. and Berger, G.W. (1988) Thermoluminescence spectra of some mineral samples relevant to thermoluminescence dating. *J. Luminescence* **39**, 123-136.

Huxtable, J., Aitken, M.J. and Bonhommet, N. (1978) Thermoluminescence dating of sediment baked by lava flows of the Chaine des Puys. *Nature* **275**, 207-209.

Juvigné, E. and Gewelt, M. (1987) La narse d'Ampoix comme téphrostratotype dans la Chaîne des Puys méridionale (France). *Bull. Ass. Fr. Etude du Quat.* **1**, 37-49.

Miallier, D., Faïn, J., Montret, M., Pilleyre, Th., Sanzelle, S. and Soumana, S. (1991) Properties of the red TL peak of quartz relevant to thermoluminescence dating. *Nucl. Tracks & Radiat. Measts.* **18(1/2)**, 89-94.

Ousmoï, M. (1989) Application de la datation par thermoluminescence au néolithique marocain. Thesis, Univ. Clermont-Ferrand, 130 p.

Pelletier, H., Delibrias, G., Labeyrie, J., Perquis, M.T. and Rudel, R. (1959) Mesure de l'âge de l'une des coulées volcaniques issues du Puy de la Vache (Puy de Dôme) par la méthode du carbone 14. *C. R. Acad. Sc. Paris* **214 D**, 2221.

Pilleyre, Th. (1991) Datation par thermoluminescence. Application à la chronologie des retombées volcaniques. Thesis. Univ. Blaise Pascal, Clermont-Ferrand, France, 170 p.

Pilleyre, Th., Montret, M., Faïn, J., Miallier, D. and Sanzelle, S. (1992) Attempts at dating ancient volcanoes using the red TL of quartz. *Quaternary Science Reviews* **11**, 13-17.

Prescott, J.R. and Stephan, L.G. (1982) The contribution of cosmic radiation to the environmental dose for thermoluminescence dating. Latitude, altitude and depth dependences. *PACT J.* **6**, 17-25.

Sanzelle, S., Erramli, H., Faïn, J. and Miallier, D. (1988) The assessment of gamma dose-rate by gamma-ray field spectrometer. *Nucl. Tracks Radiat. & Measts.* **14 (1/2)**, 209-213.

Stuiver, M., Kromer, B., Becker, B. and Ferguson, C.W. (1986) Radiocarbon age calibration back to 13,300 years BP and the ^{14}C age matching of the german oak and US bristlecone pine chronologies. *Radiocarbon* **28**, No 2B, 969-979.

Vogel, J.C. (1987) Calibration of radiocarbon dates beyond 10,000 BP. In *Chronologies in the near East* (Aurenche, O., Evin, J. and Hours, F. eds.) *BAR International Series*, 318-321.

PR Vagn Mejdahl

Suggestions for minimum requirements for reporting ESR age estimates

Rainer Grün
Radiocarbon Dating Research Unit, RSPacS, ANU, GPO Box 4, Canberra ACT 2601, Australia.

Although ESR spectrometry has been applied as a dating technique since 1975 (Ikeya 1975), there are no standards for the citation of ESR age estimates, not even any minimum requirements. It is therefore not surprising that the quality of the published procedures and analytical results for ESR age estimates vary to a great extent. In many cases ESR dating papers give little information about the estimation of the dose, D_E, or the dose rate, $\dot{D}$, that the sample received in its geological past. In order to establish ESR as a mature dating technique, it is necessary to standardize the citation procedures to some extent. It must be the aim of such a standardization, that new ESR age estimates are published with all analytical values and their errors so that the results can be reproduced/recalculated by any ESR specialist. Sufficient detail can either be given in the publication containing new age estimates or in a separate ESR data list as has been done for ^{14}C and TL age estimates in the journals *Radiocarbon* and *Ancient TL*, respectively.

Additionally, there is no agreement on the terminology used in ESR dating. The use of the acronym AD for either accumulated dose and annual dose should certainly be avoided. Additionally, it is not common practice in physics or mathematics to abbreviate a single parameter by two or more characters in sequence. It seems advantageous to adopt the terminology analogous to that of TL dating. This paper tries to outline the minimum requirements for the citation of ESR age estimates and to suggest appropriate terminology for use in ESR dating papers. It should be noted here, that the scientific committee of the TL and ESR Dating Seminars have agreed that minimum citation requirements and terminologies have to be established prior to the next meeting in Vienna 1993 and that contributions for the conference proceedings ought to follow these guidelines. Rather than being a critique of the terminology that has been forwarded by Aitken (1992) this paper intends to broaden the base for discussion.

An ESR age estimate involves basically three procedures: the determination of the radiation dose that a sample has received in the past using ESR spectrometry, the estimation of the dose rate by various different techniques and the assessment of the error. The significance of the last procedure is not often fully appreciated.

Determination of the past radiation dose, D_E

The parameter that is actually determined by ESR spectroscopy is the past radiation dose. This is the energy dose that was transferred to the mineral by ionizing radiation and the SI unit is the gray (Gy) (see e.g. The Symbols Committee of the Royal Society, 1975). This parameter is usually determined by the additive dose technique using gamma radiation from a ^{60}Co or ^{137}Cs source. The result is therefore the absorbed gamma equivalent dose which may be expressed by either D_E or $D_{E\gamma}$. The term equivalent dose, ED, is not strictly correct, since this is reserved in radiation protection for the radiation dose equivalent to humans and would have the unit sievert (Sv) (see e.g. Handbook for Chemistry and Physics).

It does not seem to be a great problem to determine a D_E value with the additional dose method (i.e. the intersect of the X-axis at I=0) either with linear or exponential fitting procedures. However, since the number of traps is limited, the trapping probability changes as traps are progressively occupied by electrons or holes and hence it is unlikely that linear fitting is the correct description of a dose response. In some cases, e.g. corals (Grün 1990), the dose response can only be described by a more complicated mathematical expression. Consequently, the mathematical model behind the curve fitting must be reported (and justified by observation).

There is no general recipe of how many different aliquots have to be measured and to which doses these aliquots should be exposed in order to obtain reliable D_E values (one might get some idea about this from the computer simulations of dose response curves (Grün & Rhodes 1991)). It is therefore necessary to cite how many aliquots and which dose steps were used for the estimation of D_E. If the ESR intensity values, I, scatter around the best fit, the D_E value is dependent on the weights, w, assigned to the intensity values. Grün & Macdonald (1989) used equal weights (w=1) whereas Berger & Huntley (1986) and Berger et al. (1987), suggested in TL studies the use of weights proportional to the inverse square of the measured intensity (i.e. $w \sim I^{-2}$, which arises from the assumption that each

ESR/TL intensity measurement has the same relative uncertainty).

Many materials such as carbonates (see Debuyst et al. 1990, 1991, Barabas 1989, Barabas et al. 1992a,b) have several ESR signals that are radiation sensitive. Hence, it has to be reported which paramagnetic centre was used for the D_E determination and how it was measured. Since D_E results derived from a particular paramagnetic centre may be dependent on the measurement conditions, such as microwave power or the resolution of the ESR spectra (see e.g. Molodkov 1988, Barabas 1989), it is necessary to cite the following parameters:

- the approximate microwave frequency, ν, in GHz or frequency band of the ESR spectrometer;
- the microwave power in mW (a dB value is related to the maximum microwave power of an ESR spectrometer and this may change from instrument to instrument). Additionally, the effective microwave power at the sample position is dependent on the sample holder, e.g. when using a double walled quartz tube, the microwave power at the sample may be twice the adjusted power;
- the scan width in tesla (T) and scan time or by combining both the scan speed;
- the modulation frequency in Hz;
- the modulation amplitude in T (often given in mTpp);
- the time constant in s.

Since it may be assumed that the spectrometer was optimally tuned before measurement, it seems unnecessary to quote the detection phase. If new materials or centres are discussed, a high resolution ESR spectrum should be shown and the g-values and line widths have to be given. It is not particularly useful to cite (instead of g-values) the microwave frequency, the magnetic field strength at the centre of the scan and the scan width so that the reader can calculate ones own g-values. If one prefers to use terms such as A, B, C.... or h_1, h_2, h_3... etc., centres, these have to be defined with g-values in a figure.

Errors in D_E estimation

There does not seem to be any agreement as to which method provides correct 1σ errors in the D_E estimation. Computer simulations of dose response curves (Grün & Rhodes 1991) imply that an error estimated on the straight line conversion of an exponential function gives errors that are significantly smaller than 1σ, whereas the jackknifing procedure (as suggested by Grün & Macdonald 1989) results in errors larger than 1σ. Additionally, the magnitude of the error seems to be dependent on the number of measurements and where these are placed in the dose response curve. More sophisticated error procedures as suggested by Berger et al. (1987) and Brumby (1992) seem to lead to consistent error estimates. It would be worthwhile to check them in computer simulations. Additionally, the radioactive source used for the artificial irradiation has a calibration error in the range of 3 to 5% as shown in the Second Interlaboratory Comparison Project on ESR dating (Barabas et al. 1992c). Although it may be assumed that dose values obtained by alanine or Fricke dosimetry are correctly converted into calcite, sulphate, quartz or apatite doses. This should be mentioned in the paper.

Determination of the dose rate, $\dot{D}$

The dose rate is defined in the Handbook of Physics and Chemistry as $\dot{D}$. If an agreement can be reached between the ESR and TL communities, that D_E or another character can be reserved for the absorbed dose, D could be used for the dose rate. This would have the advantage that one does not have to produce the dot on the D, an inconvenient procedure for standard computers. An alternative could be the German notation D' (D prime). In order to distinguish dose and dose rate parameters more clearly, in the following text $\dot{D}$ stands for dose rate. Alpha, beta and gamma dose rates may be denoted by subscripts ($\dot{D}_{\alpha,\beta,\gamma}$).

Both notations, $\dot{D}$ and D', suggest an instantaneous dose rate. However, most Quaternary samples show radioactive disequilibria and the cited dose rates are usually average dose rates over the calculated age of the sample. In these cases a notation implying a derivative is not strictly correct.

The dose rate is determined by the analysis of the radioactivity in the sample and its surrounding. This can either be done by measuring the total alpha or beta activities or by the analysis of the radioactive elements. The latter values have to be converted into dose rates by using published tables. It must be noted that these tables (e.g. Bell 1976, 1977, 1979, Hennig & Grün 1983, Nambi & Aitken 1986 or Berger 1988) have different conversion factors and the respective publication has to be cited.

It should be noted which analytical techniques were used and which parameters were assumed or taken from literature (e.g. water content, $^{234}U/^{238}U$ ratio or α-efficiency). Depending on the symmetry, a different set of parameters has to be measured.

In general the following parameters have to be reported:

Cosmic dose rate, $\dot{D}_{cos}$: the cosmic dose rate can either be measured with calibrated gamma spectrometers or derived from the thickness of the overlying sediment (Prescott & Hutton 1988).

Alpha efficiency: this value is very rarely measured in ESR dating studies and the discussion whether to use a k-value (Zimmerman 1972), an a-value system (see Aitken 1985) or a b-value system (Bowman & Huntley 1984, Huntley et al. 1988) etc. is rather academic.

However, it has to be reported which value has been used or why a particular α-efficiency was assumed (e.g. taken from literature). The following measured α-efficiencies have been reported: g=2.0007 in corals: 0.06±0.01 (Radtke & Grün 1988) and 0.05±0.01 (Grün et al. 1992); g=2.0007 in foraminifera: 0.079±0.008; g=2.0036 in foraminifera: 0.093±0.009 (Mudelsee 1990, Mudelsee et al. 1992); g=2.0007 in speleothems: 0.052±0.026 (Lyons & Brennan 1991); g=2.0018 in tooth enamel: 0.15 (Grün 1985; DeCanniere et al. 1986).

Radioactive disequilibria: Most samples that are investigated in dating studies, such as secondary carbonates or tooth enamel, show disequilibria in the U-decay chains. This can be mathematically considered (for the respective formulae see e.g. Grün, 1989). Additionally, many samples show $^{234}U/^{238}U$ ratios of greater than unity. This is rarely measured. In some cases it may be derived from U-series measurements (e.g. secondary carbonates in sea water have the initial $^{234}U/^{238}U$ ratio of 1.14±0.01: Chen et al. (1986)). An assumed value of 1.2±0.2 should normally not lead to large systematic errors. However, one has to be aware that in some areas strong $^{234}U/^{238}U$ disequilibria may occur (see e.g. multiple dating study in Egypt by Wendorf et al. 1990: $^{234}U/^{238}U << 1$!).

External gamma dose rate, $\dot{D}_\gamma$ The external gamma dose rate cannot be assumed. It can be measured by a portable gamma spectrometer or TL dosimeters. Both techniques have a calibration error which is usually in the range of 5 to 7%. Sometimes $\dot{D}_\gamma$ is derived from the analysis of the radioactive elements in the sediment. However, this approach can only be used for very homogeneous sediments (such as loess) and may lead in other cases to significant errors. The analytical technique has to be mentioned along with any further information about radioactive disequilibria.

The dose rate of a sample is strongly dependent on its size, because of the different ranges of alpha, beta and gamma rays. One can distinguish four principal cases:

1) Infinitely thick (homogeneous) samples (>60 cm, e.g. corals): the total dose rate (except cosmic) is generated by the sample. From the analysis of the radioactive elements (namely U, Th, K), the α, β, and γ dose rates are derived. Corals, for example, contain only small amounts of Th and K. It should be reported if these have been neglected. It is always worthwhile to check whether the calculated gamma dose rate can be confirmed by measurements. The following parameters have to be determined:

- U, Th, K and water contents of the sample.
- U-series disequilibria (i.e. $^{234}U/^{238}U$, $^{230}Th/^{234}U$ and $^{231}Pa/^{235}U$)
- α-efficiency
- cosmic dose rate

2) Thick samples (≥ 6mm and < 60 cm, e.g. mollusc shells): the outer 2 mm can be removed and the sample receives only external $\dot{D}_\gamma$. If the sample is thicker than a few cm, the measured external gamma dose rate is attenuated by the sample and this effect has to be considered. For preliminary results on grains see Mejdahl (1983), for layers see Aitken et al. (1985). The following parameters have to be determined:

- U, Th, K, and water contents of the sample
- U-series disequilibria (i.e. $^{234}U/^{238}U$, $^{230}Th/^{234}U$ and $^{231}Pa/^{235}U$)
- α-efficiency
- gamma self-irradiation
- external gamma dose rate
- gamma attenuation
- cosmic dose rate

3) Medium samples (>200 μm and <6mm, e.g. tooth enamel): it is possible to remove about 20 μm to eliminate the external alpha dose rate, however, the sample receives a portion of the external beta dose rate. Beta particles are attenuated and this attenuation has to be considered in the dose rate calculation. For beta attenuation in grains see Mejdahl (1979), for thin layers see Grün (1986). The beta emitters in the immediate surroundings of the sample (generally within 2mm) have to be determined along with the following values:

- U, Th, K and water contents of the sample
- U-series disequilibria (i.e. $^{234}U/^{238}U$, $^{230}Th/^{234}U$ and $^{231}Pa/^{235}U$)
- α-efficiency
- beta self-irradiation
- external beta dose rate (U, Th, K and water of the immediate surroundings)
- beta attenuation
- external gamma dose rate
- cosmic dose rate

4) Small samples (< 200 μm, e.g. foraminifera): it may not be possible to remove the volume that has received an external alpha dosage. Alpha and beta attenuation has to be considered. For the alpha attenuation of grains see Bell (1980), for layers see Aitken (1987) and Grün (1987). The following results have to be reported:

- U, Th, K and water contents of the sample
- U-series disequilibria (i.e. $^{234}U/^{238}U$, $^{230}Th/^{234}U$ and $^{231}Pa/^{235}U$)
- α-efficiency
- alpha self-irradiation
- beta self-irradiation
- external beta and alpha dose rate (U, Th, K and water of the immediate surroundings)
- alpha attenuation
- beta attenuation
- external gamma dose rate
- cosmic dose rate

For very small samples, such as foraminifera, it may be assumed that the alpha dose rate corresponds more or less to the infinite matrix dose (including the samples, see Mudelsee et al. 1992). In other cases, e.g. small quartz grains, this may not be valid.

These lists show the parameters that have to be cited in principle, but they may not necessarily apply to each dating approach. For example, one can estimate the external beta dose rate by (i) measuring U, Th, K of the immediate surrounding using neutron activation analysis and the water content or by (ii) measuring the beta activity of the immediate surroundings with, for example, a thick source beta counter, TSBC (see Sanderson 1988) and the water content. Each approach has its own merits. In order to measure attenuation factors, the dimensions of the sample have to be measured and account taken of any layers that have been removed during sample preparation.

Errors in $\dot{D}$ determination

The exact error calculation for the total dose rate is rather complex. It is easily possible to estimate the errors for the analytical techniques, since they are usually based on counting statistics. Other sources of errors, however, are rather difficult to estimate, such as the average water content in the sediment. One solution seems to be measurement of the present day water content and to assume a relative large error (e.g. 10±5 % by weight). Berger (1988) suggests to measure the present saturation water content and derive the average water concentration and error from the in situ water content and site information. As long as no U-series disequilibria are involved in the age calculation, Appendix B of Aitken (1985) deals extensively with the error calculation of TL (and hence ESR) age estimates. Since the formulae for U-series disequilibria are rather lengthy (see e.g. Grün 1989) it seems technically simpler to estimate the influence of each uncertainty on the calculation of the total dose rate. This approach is valid as long as the error is approximately linear and corresponds to a numerical derivation. The total error results from the square root of the sum of squares of the individual error contributions.

Some parameters whose occurrence is rather uncertain, such as Rn loss or U-series disequilibria in sediments, should only be included in the estimation of the dose rate if they have actually been measured. Otherwise, readers may get the impression that the introduction of those processes was used to arrive at a convenient result! In practice, most of the general information can be given in the foot note of a table.

I wish to emphasise that at present the points outlined above may be regarded as a basis for discussion. I hope that many further suggestions will be made to improve these suggestions and to establish firm requirements for the citations of ESR age estimates prior to the next International Seminar on TL and ESR Dating in Vienna in 1993.

Acknowledgements

I thank M.J. Aitken, G. W. Berger, S. Brumby, R. Debuyst, F. Dejehet, M. Ikeya, V. Mejdahl, U. Radtke, E.J. Rhodes, J. Rink, H.P. Schwarcz, A. Skinner, R. Walther, A. Wieser and A.G. Wintle for many helpful comments.

References

Aitken, M.J. (1985) *Thermoluminescence dating*. London, New York, Academic Press, London, 359 pp.

Aitken, M.J. (1987) Alpha dose to a thin layer. *Ancient TL* **5(3)**, 13

Aitken, M.J. (1992) Symbols in TL and optical dating, provisional list. *Ancient TL* **10**, 1516.

Aitken, M.J., Clark, P.A., Gaffney, C.F. & Lovborg, L. (1985) Beta and gamma gradients. *Nucl. Tracks* **10**, 647-653.

Barabas, M. (1989) ESR Datierung von Karbonaten, Grundlagen, Systematic, Anwendun gen. Inaugural Dissertation, Universität Heidelberg, 161 pp.

Barabas, M., Mudelsee, M., Walther, R. & Mangini, A. (1992a) Dose response and thermal behaviour of the ESR signal at g=2.0006 in carbonates. *Quaternary Science Reviews* **11**, 165-171.

Barabas, M., Mudelsee, M., Bach, A., Walther, R. & Mangini, A. (1992b) General properties of the paramagnetic centre at g=2.0006 in carbonates. *Quaternary Science Reviews* **11**, 173-179.

Barabas, M., Walther, R., Wieser, A., Radtke, U. & Grün, R. (1992c) Second Inter laboratory Comparison Project on ESR Dating. *Applied Radiation and Isotopes* (in press).

Bell, W.T. (1976) The assessment of the radiation dose rate for thermoluminescence dating. *Archaeometry* **18**, 107-111.

Bell, W.T. (1977) Thermoluminescence dating, revised doserate data. *Archaeometry* **19**, 99-100.

Bell, W.T. (1979) Thermoluminescence dating: radiation dose-rate data. *Archaeometry* **21**, 243- 245.

Bell, W.T. (1980) Alpha dose attenuation in quartz grains for thermoluminescence dating. *Ancient TL* **12**, 48.

Berger, G.W. (1988) Dating Quaternary events by luminescence. In, *Dating Quaternary Sediments* (Ed. D.J. Easterbrook) Spec. Pap. Geol. Soc. Am. 227, 13-50.

Berger, G.W. & Huntley, D.J. (1986) Linear regression of TL data. *Ancient TL* **4**, 26-29.

Berger, G.W., Lockhart, R.A. & Kuo, J. (1987) Regression and error analysis applied to the dose response curve in thermoluminescence dating. *Nucl. Tracks Radiat. Meas.* **13**, 177-184.

Bowman, S.G.E. & Huntley, D.J. (1984) A new proposal for the expression of alpha efficiency in TL dating. *Ancient TL* **2**, 68.

Brumby, S. (1992) Regression analysis of ESR/TL dose response data. *Nucl. Tracks Radiat. Meas.* (in press).

Chen, J.H., Edwards, R.L. & Wasserburg, G.J. (1986) ^{238}U, ^{234}U and ^{232}Th in seawater. *Earth and Planetary Science Letters* **80**, 241-251.

Debuyst, R., Bidiamambu, M. & Dejehet, F. (1990) Diverse CO_2^- radicals in γ- and α- irradiated synthetic calcite. *Bull. Soc. Chim. Belg.* **99**, 535-541.

Debuyst, R., Bidiamambu, M. & Dejehet, F. (1991) An EPR study of γ– and α- irradiated synthetic powdered calcite labelled with carbon 13. *Nucl. Tracks Radiat. Meas.* **18**, 193-201.

DeCanniere, P., Debuyst, R., Dejehet, F., Apers, D. & Grün, R. (1986) ESR dating, A study of ^{210}Po coated geological and synthetic samples. *Nucl. Tracks* **11**, 211-220.

Grün, R. (1985) Beiträge zur ESR Datierung. Sonderveröffentlichungen des Geologischen Instituts der Universität zu Köln, **59**, 11-57.

Grün, R. (1986) Beta dose attenuation in thin layers. *Ancient TL* **4**, 18.

Grün, R. (1987) Alpha dose attenuation in thin layers. *Ancient TL* **5**(3), 6-8.

Grün, R. (1989) Electron spin resonance (ESR) dating. *Quaternary International* **1**, 65-109.

Grün, R. (1990) Dose response of the paramagnetic centre at g=2.0007 in corals. *Ancient TL* **8**(3), 20-22.

Grün, R. & Macdonald, P.D.M. (1989) Nonlinear fitting of TL/ESR dose response curves. *Applied Radiation and Isotopes* **40**(10-12), 1077-1080.

Grün, R. & Rhodes, E.J. (1991) On the selection of dose points for ESR/TL dose response curves. *Ancient TL* **9**, 40-46.

Grün, R., Radtke, U. Omura, A. (1992) ESR and U-series analyses on corals from Huon Peninsula, New Guinea. *Quaternary Science Reviews* **11**, 197-202.

Grün, R., Schwarcz, H.P. & Zymela, S. (1987) ESR dating of tooth enamel. *Canadian J. of Earth Sciences* **24**, 1022-1037.

Hennig, G.J. & Grün, R. (1983) ESR dating in Quaternary geology. *Quaternary Science Reviews* **2**, 157-238.

Huntley, D.J., Berger, G.W. & Bowman, S.G.E. (1988) Thermoluminescence responses to alpha and beta irradiations and age determination when the high dose response is nonlinear. *Radiation Effects* **105**, 279-284.

Ikeya, M. (1975) Dating a stalactite by electron paramagnetic resonance. *Nature* **255**, 48-50.

Ikeya, M. (1982) A model of linear uranium accumulation for ESR age of Heidelberg (Mauer) and Tautavel bones. *Japanese J. of Applied Physics* **21**, L690-L692.

Lyons, R.G. & Brennan, B.J. (1991) Alpha/gamma effectiveness ratios of calcite speleothem. *Nucl. Tracks Radiat. Meas.* **18**, 223-227.

Mejdahl, V. (1979) Thermoluminescence dating, beta dose attenuation in quartz grains. *Archaeometry* **21**, 61-72.

Mejdahl, V. (1983) Feldspar inclusion dating of ceramics and burnt stones. *PACT* **9**, 351-364.

Molodkov, A. (1988) ESR dating of Quaternary mollusc shells recent advances. *Quaternary Science Reviews* **7**, 477-484.

Mudelsee, M. (1990) ESR-Datierung karbonathaltiger, quartärer Tiefseesedimente. Unpublished Diplomarbeit. Universität Heidelberg, 73pp.

Mudelsee, M., Barabas, M. & Mangini, A. (1992) ESR dating of the Quaternary deep sea sediment core RC17177. *Quaternary Science Reviews* **11**, 181-189.

Nambi, K.S.V. & Aitken, M.J. (1986) Annual dose conversion factors for TL and ESR dating. *Archaeometry* **28**, 202-205.

Prescott, J.R. & Hutton, J.T. (1988) Cosmic ray and gamma ray dosimetry for TL and ESR. *Nucl. Tracks Radiat. Meas.* **14**(1/2), 223-227.

Radtke, U. & Grün, R. (1988) ESR dating of corals. *Quaternary Science Reviews* **7**, 465-470.

Sanderson, D.C.W. (1988) Thick source beta counting (TSBC), a rapid method for measuring beta dose-rates. *Nucl. Tracks Radiat. Meas.* **14**(1/2), 203-207.

The Symbols Committee of the Royal Society (1975) Quantities, Units, and Symbols. 2nd Edition, 54p., The Royal Society, London.

Wendorf, F., Close, A.E., Schild, R., Gautier, A., Schwarcz, H.P., Miller, G.H., Kowalski, K., Krolik, H., Bluszcz, A., Robins, D. & Grün, R. (1990) Le Dernier interglaciaire dans le Sahara oriental. *L'Anthropologie* **94**, 361-391.

Zimmerman, D.W. (1972) Relative thermoluminescence effects of alpha and beta radiation. *Radiation Effects* **14**, 81-92.

PR. Members of the Editorial Board and H.P. Schwarcz

Estimation of accumulated dose and its uncertainties: potential pitfalls in curve fitting

R.G. Lyons†, B.J. Brennan‡ and P.L. Hosking*
†Department of Geography and Oceanography, Australian Defence Force Academy, University College, University of New South Wales, ACT 2600, Australia.
Departments of ‡Physics and *Geography, University of Auckland, New Zealand.

Introduction

The problems of curve fitting in general, and determining uncertainties in AD in particular, are recognized as being of significance in ESR and TL dating, as evidenced by the recurrence of the topic over recent years (Rendell, 1985; Berger and Huntley, 1986; Franklin, 1986; Berger *et al.*, 1987; Lyons, 1988; Grün and MacDonald, 1989; Berger, 1990). This discussion aims to clarify for non-statisticians some of the issues which have arisen, to focus attention on potential problems, and, moreover, to encourage available techniques to be applied critically rather than naïvely. It also draws attention to some of the recent developments in statistical theory.

Relevant statistical concepts

For simplicity, the following examples use linear fitting to demonstrate some important concepts, but the concepts do, of course, apply equally to any other curve.

(i) Leverage.

Points will have a greater or less effect on the fitted curve depending on their position on the growth curve. As noted by Franklin (1986), end points are more influential in determining the position of the line, particularly for extrapolation, than more central points (fig. 1; table 1). Depending on whether the extreme points are "good" or "bad", leverage is either a desirable or undesirable property: a 'good' point is one which has both a low uncertainty (high precision) and is a true representation of the physical reality (high accuracy), while a 'bad' point has a low precision and/or accuracy (see also fig. 2 and the discussion below on outliers).

Confidence in the accuracy of the natural non-irradiated signal intensity should be well-founded. The high leverage exerted by its position can therefore be justified provided it also has a high precision. The leverage of low dose data points is particularly important if the Y-values have equal variance. It is then highly desirable to reduce the variance of the natural signal by repeating the measurement and using the average value in the curve fitting procedure (see later, on replication of data points). If, on the other hand, the data are believed to have a constant *percent* error, then the low dose points are "good" and their leverage is a desirable property; the data set will be enhanced by increasing the number of low dose points.

It is unfortunately true that high dose points are generally more suspect as being further removed from the environmental situation; repeated measurements will improve their precision but not their intrinsic accuracy. Even where the high dose points have a small uncertainty, a serious problem arises if the model fitted does not fully represent the underlying physical behaviour because they will be "bad" points with respect to the fitted model, and will exercise undesirable leverage (see fig. 4).

Leverage by high dose points can be reduced by two means:

(a) inverse weighting to redress the implicit heavy weighting they receive by virtue of their position. (Note that this inverse weighting is in addition to any required by the unequal variances, or precision, of the data.) Carrying out a log-transformation to the data in order to fit a linear instead of an exponential model, as in Apers et al., (1981), effectively applies a weighting of $1/Y_i$ and is desirable, rather than to be avoided as suggested in Grün and MacDonald (1989). Note Grün and Rhodes (1991) in estimating the optimal distribution of data points for the dose response curve fitted their models to log-transformed data, thereby reducing the effect of leverage by the high dose points. Their conclusions are therefore only valid for similarly weighted data. If curve fitting procedures are used which do not inversely weight the data, leverage by high dose points will be excessive and lower maximum doses should be used.

(b) provision of an adequate range of intermediate points. This will reduce the effect of leverage of high dose points where the model is adequate but the precision of the data poor. Intermediate points also indicate whether the model is appropriate (see below).

This discussion should not be taken to imply that high dose points are not of value. On the contrary, they are essential to evaluate the appropriateness of the model and to provide robustness to the extrapolation.

(ii) Influence.

The poor quality of a high leverage point is often masked because it tends to draw the fitted curve towards itself (see figs 1 & 2). This phenomenon is known as *influence*, and has become a major focus for statisticians

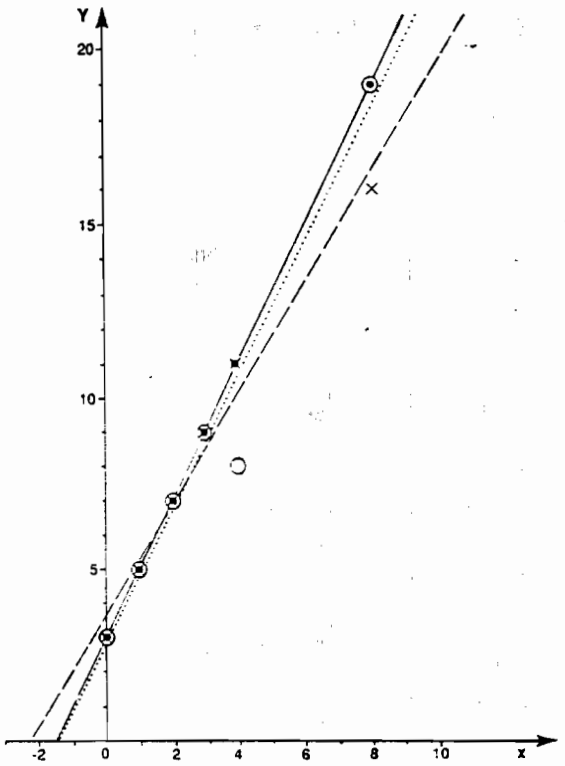


Figure 1.
Leverage in regression curve fitting. The solid line is the model line Y=2x+3, fitted to the infilled circles. Where an extreme point is "in error" (dashed line, fitted to the crosses), the regression line shows pronounced leverage with both AD and slope being significantly affected. Because an extreme point draws the line towards itself (influence), its poor quality is not necessarily obvious and it may not have a high residual. When a central point is in error, the effect on AD and slope is reduced, involving a sub-parallel shift instead of leverage (dotted line fitted to open circles). The effects of the same absolute or percentage change on the lowest point is given in table 1.

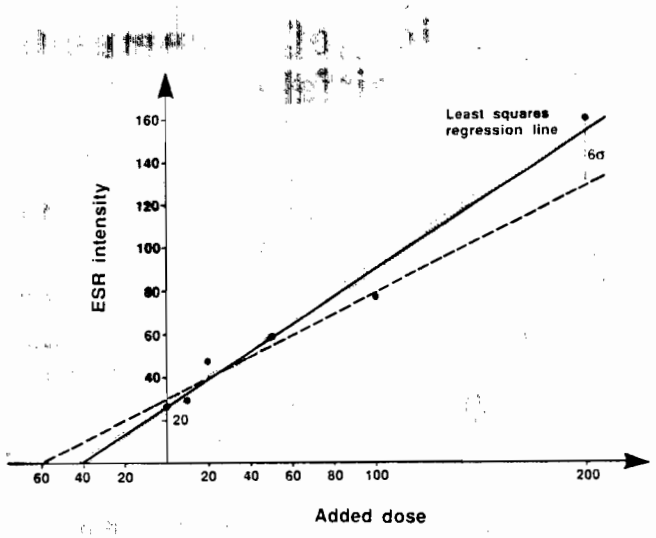


Figure 2.
Outliers produced by jackknifing reflect the quality of the associated jackknifed (omitted) point: when a poor quality high leverage point is omitted a high quality outlier is produced and vice versa. The data points have been produced for the model line (dashed line, y=30+0.5x, AD=60) with randomly generated errors, s=5, assigned to the Y values. To demonstrate the effect of leverage on a scale which can be graphically shown, the high leverage point has been deliberately assigned a value which places it 6σ from the model line (in fact, if such a point arose experimentally and its deviation from the line were known, it would no longer be regarded as a member of the data set). However, note that the regression line to all data points (solid line) passes close to the poor quality point by reason of its influence; the quality of a point cannot be determined by visual examination nor by residuals (table 2) but must be independently known. Omitting the poor quality high leverage point in the jackknifing procedure produces an outlier which lies too close to the model line to be distinguished on this scale. The dotted lines are the jackknife fits when the other points are omitted and, although closely grouped, give poor estimates of the model parameters.

in the last decade (Cook and Weisberg, 1982). An important consequence of influence is the difficulty of determining whether high leverage points are 'good' and hence whether the associated estimates are biased.

A plot of the fitted curve together with the data points serves only as a first visual test of the goodness of fit. Because of influence, an apparently good fit is, in itself, no guarantee of the quality of any particular data point nor that the best curve has been fitted. Mathematically, this means that a high or low residual cannot be used to identify a poor or high quality data point, respectively (table 2). Any attempt to 'improve' a data set by eliminating high residual point(s) is at best misguided and potentially disastrous. Thus, tests such as Chauvenet's test for identifying outlying data points (Misaqi, 1975) are not sufficient grounds for rejecting high leverage data points, as these tests essentially depend on some form of evaluation of residuals. The quality of a high leverage data point should be assessed independently of its relationship with the fitted curve, for example, by its variance (which can then be included as a weighting factor) or other experimental evidence.

Assessing model aptitude

The choice of model is of importance in any application involving extrapolation. Models which are satisfactory for interpolation (such as high order polynomials) may be completely unsatisfactory for extrapolation. It is not intended here to discuss the appropriateness of any particular model, but simply to identify some relevant statistical procedures and pitfalls. Note that Readhead (1984, 1988) and Prescott (1983) propose a regenerative method of estimating AD in TL dating in which the curve fitting utilizes interpolation rather than

extrapolation; it is therefore much less vulnerable to incorrect model choice.

The phenomenon of leverage can be used informally to test the appropriateness of a model. If the AD estimated by the model shows a systematic trend as additional higher value points are progressively added to the calculation, then the model is suspect, and the effect of leverage may be extremely serious. (This technique is used by Grün (1990) to demonstrate the inappropriateness of the simple saturating exponential model to describe the growth curve of the g=2.0007 centre in corals.)

Just as a low residual at a single point is not necessarily an indicator of a 'good' point, neither is a low residual sum of squares (SSE) an indicator of a 'good' model, as is sometimes assumed. Introducing more parameters into a model will *always* reduce the residual sum of squares, because SSE is an estimate of the variation in the data left unexplained by the model - the more complex the model, the closer the fit and the "better" the correspondence. A better indicator of the fit of a model is the mean sum of squares (MSE) , as this allows both for the number of parameters estimated and the number of data points:

$$MSE = SSE/(n-p)$$

where,

SSE = residual sum of squares (error sum of squares)
n = no of data points
p = no of parameters fitted (incl. intercept term)

An equally useful indicator is the root mean square error, or the standard error of estimate (RMS), which is simply the square root of MSE but has the advantage that it is in the same units as the dependent variable.

The most common method to assess model adequacy, which takes into account all of the above, is to consider the ratio of MSR/MSE, where:

$$MSR = SSR/(p-1)$$

where,

SSR = regression sum of squares
MSE = mean sum of squares, as above.

This ratio has an F distribution, F(p-1, n-p). Thus the F-statistic (or the associated probability of rejection) is a useful measure for comparison of models. The Student's t-test is the corresponding test for the linear case.

Estimation of uncertainty in AD

The estimate of uncertainty in AD is a particularly thorny problem. Because the data for experimentally determined additive growth curves may not satisfy some of the basic statistical assumptions, caution is needed in applying standard statistical techniques.

Four methods have been suggested in the ESR and TL literature: regressing x on Y (Rendell, 1985); quasi-likelihood techniques (Berger *et al.*, 1987); constrained linear regression (Lyons, 1988); and jackknifed regression (Grün and MacDonald, 1989). Which method is considered the most appropriate will depend on the nature of the data.

The first of these methods, regressing x on Y, contravenes the basic assumption that the error is associated only with the dependent variable and that the independent variable is error-free. As a result the 'regression may be so greatly changed as to be meaningless' (Williams, cited in Rendell, 1985). While quasi-likelihood methods are statistically established techniques and their use is to be recommended where possible, it is also true that they are generally 'computationally demanding' (Seber, 1977) and may place unrealistic demands on both the quantity and quality of the data, requiring 'at least 15 points (Berger *et al.*, 1987), each with an uncertainty of <5%): meeting these requirements must be carefully assessed. Constrained linear regression (Lyons, 1988), which forces the regression line to pass through the natural signal intensity, does not have these restrictions on data quality and quantity but requires the natural signal to be extremely well-defined. It is only applicable to a linear model and thus its usefulness is limited. Re-sampling techniques such as jackknifing (see below), may be applied to any model to obtain uncertainty estimates: their validity with respect to regression is, however, the topic of debate because two basic statistical assumptions may not be satisfied by regression data:

(i) independent random data points. This is untrue for additive growth curve data, where the independent variable (added irradiation) is not randomly selected.
(ii) equal variance, i.e. the total *absolute* experimental uncertainty is the same for each point regardless of its magnitude. While this assumption is supported by some (Franklin, 1986; Scott and Sanderson, 1988), others would regard a constant *percentage* error for each data point as more reasonable (Berger, 1984; 1987).

Re-sampling techniques

Re-sampling techniques require constructing subsets of data from the initial complete data set and noting how the parameter of interest varies with the different subsets. For example, in 'simple delete-one jackknifing', as used in the program FITT supplied by Grün and MacDonald, the subsets of the data are obtained by omitting each data point in turn from the total set. The uncertainty in AD is calculated from the resulting distribution of AD_i. (While this description is conceptually apt, it should be noted that the mathematical theory and execution is more complex.)

Reasonably comprehensible descriptions of re-sampling techniques for the general user have been written by

Miller (1974) and Efron (1982). More sophisticated treatments can be found in journals such as the *Annals of Statistics* and the *Journal of the American Statistical Association*. In particular, Vol 14 No. 4, *Annals of Statistics*, contains a substantive invited paper by Wu (1986), including a comprehensive reference list, and a series of related discussion papers.

Re-sampling techniques such as jackknifing have wide application in experimental analysis. However, most of the theoretical work on re-sampling is based on the two assumptions outlined above and thus is not necessarily valid for regression (Wu, 1986; Carroll and Rupert, 1986; Hinkley, 1977). The application of resampling techniques to regression data is currently the subject of much debate in statistical research (see *Annals of Statistics*, Vol.14, No. 4) and 'more work is needed and is being done in this area' (Shao, 1986). Points of concern are potential bias in the parameter estimates, possibly unduly inflated (conservative) or skew estimates of the associated uncertainties (Hall, 1986), underestimation of uncertainties (Duncan, 1978, cited in Seber and Wild, 1989), and the disproportionate effect of outlying data points on the results. Re-sampling techniques can also be computationally demanding, and methods are being developed to bring the computing requirements within manageable limits for large data sets (not a problem with TL and ESR applications!) It should be remembered that statisticians almost always work with very large data sets, which tends to minimize the problems mentioned above. For the small data sets used in AD determinations, the computational demands are modest but these problems may be of much greater practical significance (Hinkley, 1977).

Developments in re-sampling techniques for estimating uncertainties in regression applications are significant. Wu (1986) suggests the use of a 'weighted sharpened delete-n jackknife' to cope with some of the problems raised by the the nature of regression data and non-satisfaction of the above assumptions. In this method, instead of only one point being omitted to form each

Table 1.
Leverage due to position: effects on AD are much greater for extreme points. Model line is Y = 2x + 3. The second set of figures for the central and bottom points have the same percentage deviation from the true Y value as the top point.

	"Bad" pt	Change in y	y-intercept	Slope	AD	% change in AD
model	none	0	3	2	1.5	
top	(8, 16)	-3	3.63	1.63	2.23	+49
middle	(4, 8)	-3	2.73	1.93	1.42	-5
	(4,9.26)	16%	2.84	1.96	1.45	-3
bottom	(0, 0)	-3	1.83	2.23	0.82	-45
	(0,2.53)	16%	2.82	2.04	1.38	-8

Table 2.
Residuals are poor indicators of data quality because of influence of extreme points. The model and data are as in fig. 2. AD_i are the jackknife estimates omitting the ith point. AD* is calculated from the average slope and the average Y-intercept from the jackknifed slope and intercept estimates. a) extreme high leverage point , poor quality but with a low residual, AD_{200} is an outlier and the closest estimate of AD, AD* is a very poor estimate of AD b) extreme point is very high quality (lies on the model line), AD_{200} is an outlier and the worst estimate of AD. Note that, in the absence of information on the quality of a high leverage point, the best estimate of AD will be obtained by fitting the data to all the points, suitably weighted to allow for leverage, as discussed in the section on the Application of the delete-one jackknife. Note also that the jackknife estimates of the uncertainty in AD given here, are substantially higher than those which would be calculated by the standard formula for independent data set because the resampled data sets used in jackknifing are not independent.

x	0	10	20	50	100	200
e_i	-3.5	-5.7	7.5	4.0	-2.8	30.0
residual	0.66	-2.98	8.79	0.99	-12.97	5.50
AD_i	39.7	42.5	34.6	39.8	41.1	58.9
	AD_{model} = 60.0	AD_{all} = 40.2		AD_{av} = 42.7 ± 17.0		AD* = 35.3
e_i	-3.5	-5.7	7.5	4.0	-2.8	0
residual	3.21	-6.08	2.54	3.52	5.85	-1.99
AD_i	55.7	66.2	57.3	61.4	59.4	51.7
	AD_{model} = 60.0	AD_{all} = 59.6		AD_{av} = 58.7 ± 10.1		AD* = 58.5

subset as in the delete-one jackknife, several (n) points selected at random are omitted. This increases the independence of the subsets and the estimated experimental uncertainty is then largely independent of the underlying theoretical distributions. It should also be readily extendable to non-linear situations (Rao and Prasad, 1986). The weighted sharpened delete-n jackknife is conceptually equivalent to the common practice in physics and chemistry, where the uncertainty of the estimate of a parameter is calculated from repeated determinations (i.e. from a number of independent data sets) in accordance with the central limit theorem.

Although the relative merits of the various methods of jackknifing and other re-sampling techniques are still a matter for discussion, progress is substantial and very promising (Shao and Wu, 1989).

Application of the delete-one jackknife

Meanwhile, even though the application of the simple delete-one jackknife to regression problems is not universally accepted by the theoreticians, it may well be argued that 'something is better than nothing'. Certainly, the availability and convenience of the FITT program for simple delete-one jackknifing, supplied by Grün and MacDonald, is a strong argument for its use. However, bearing in mind the concepts discussed above, and the emphatic comment by Wu (1986) that 'routine and blind use of any re-sampling technique is problematic', several points arise:

1) Treatment of outliers.

The problem of leverage may lead to outliers, or values which are substantially different to most of the values, in the set of AD_i generated and hence to highly dubious estimates of both AD and the uncertainty in AD. Depending on the quality of the point omitted to obtain the outlying estimate, the outlier will represent either the best (poor quality point omitted) or the worst (high quality point omitted) estimate of AD (fig. 2 & table 2). It is essential to evaluate the quality of the outlier-producing point independently (NB it cannot be assessed simply from examination of the residual) in order to interpret the jackknife data correctly. A point must not be discarded simply because it produces an outlier AD estimate unless its exclusion is justified on other grounds. The outlying estimate may be the closest estimate of AD if it arises from excluding a poor quality high leverage point.

The above comments are not an objection to identifying and, possibly, eliminating outliers, nor to the sometimes large estimates of uncertainty that may arise if the data set itself contains outliers, as was mistakenly interpreted by Grün (1990). Rather, we are concerned that outliers in the resulting estimates of AD and curve parameters should be closely scrutinized for their significance.

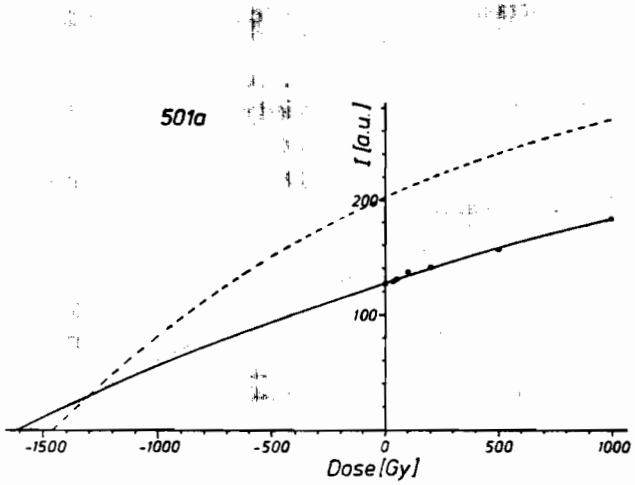


Figure 3.
Estimation of AD from jackknife estimates and the effect of duplicating points (from Grün and MacDonald, 1989; the labelling of the 3 lines has been corrected, with permission). The solid curve is the best fit to all points. The parameters obtained by averaging the estimates from jackknifing for each curve parameter may give rise to a curve with little correspondence to the data points (dashed curve). The dotted curve is the result of averaging the estimates from jackknifing after adding one additional (assumed) point at 1 kGy. Contrary to Grün and MacDonald's statement, this "improvement" is not due to the "(better) choice of radiation steps", but to the forced inclusion of the 1 kGy value, by duplication, in all the jackknifed estimates i.e. by the elimination of the corresponding outlier.

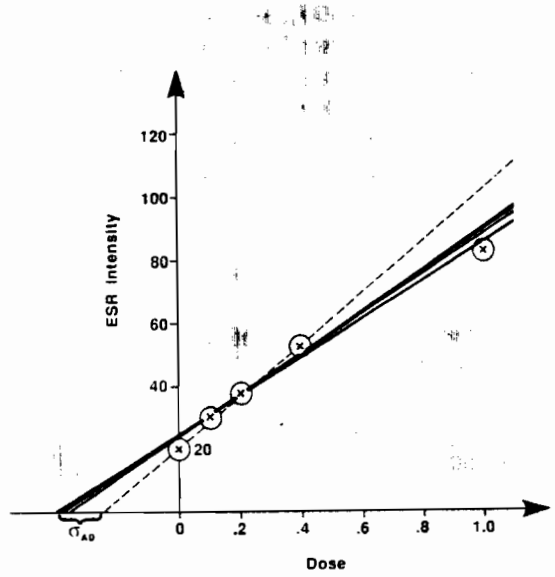


Figure 4.
Model aptitude cannot be tested by the effect of duplication on jackknife uncertainty estimates. The data is saturating exponential but the fitted model is linear; the extreme point gives rise to an outlier. Duplicating this point (or including another high value point close to it) effectively prevents its elimination and will reduce s_{AD} without improving the fit; s_{AD} from jackknifing cannot be used to determine model fit.

The vulnerability of the jack-knife technique to outliers in the parameter estimates and the effective weighting it gives to extreme points when extrapolation is being used, mean that it must be applied with caution, and due consideration given to appropriate weighting factors. Its use should also be restricted to high quality data if the resulting estimates of uncertainty are to be realistic.

2) Estimation of AD.

The best estimate of the curve parameters should be calculated by fitting all the data. It should *not* be derived by numerically averaging the estimates of the curve parameters obtained in the jackknifing procedure (e.g average Y-intercept/average slope for a linear model), as these variables are not independent. Nor should it be obtained from the average of the AD_i from jackknifing (table 2) because these AD_i are not of equal value due to leverage effects. For example, a poor quality high leverage point may lead to severe bias in the estimate of AD, because it will give an AD significantly different from the true estimate. This will have a disproportionate effect on the average value of AD, both because of the size of the discrepancy and because it is inherently a poor quality estimate and therefore should not receive equal weighting in the averaging process (Carroll and Rupert, 1986).

If either of these averaging methods is used, the AD estimate may not only be biased but the resulting curve an obviously poor fit, as shown in fig. 3 from Grün and MacDonald (1989). More sophisticated averaging techniques have been proposed to deal with these difficulties: Hinkley (1977) and Siminoff and Tsai (1986) propose using 'pseudovalues' to downweight high leverage data when calculating the jackknifed estimates. These effectively weight the data to allow for leverage effects when each jackknifed estimate is calculated, thus reducing the bias. The problem is more complex when the data points have unequal variances (Freedman, 1986), as eliminating a high variance point yields an inherently 'better' estimate than one where a low variance point has been omitted, but should be amenable to the same technique of pseudovalues using the Fisher information matrix (Wu, 1986).

3) Replication and distribution of data points.

The magnitude of the uncertainty estimated by jackknifing depends heavily on the distribution of the data points. Replication of points will reduce the estimate markedly, and spuriously, as it effectively prevents the elimination of the replicated data point in the jackknifing procedure (fig. 4). Instead of including any replications individually when jackknifing is used to estimate uncertainty in AD, they should be averaged and the resulting single point given an increased weighting because of its reduced uncertainty. Note that if other techniques, such as quasi-likelihood methods, are used, averaging of replicated points is not required. Indeed, except when resampling techniques are used, maximum value is obtained from the data if all points are included in the curve fitting process individually

Any apparent reduction in AD uncertainty as estimated due to duplicating data will be largely spurious and the AD estimate itself will be biased. The 'demonstration' in Grün and MacDonald (1989) of dramatically improved uncertainty limits and 'better' curve fitting as a result of duplication is incorrect and seriously misleading (fig. 3). On the contrary, the example demonstrates, not that there has been a 'wrong choice of radiation steps', but that a good distribution of points over the whole range of the independent variable is vital to enable the validity of the model to be assessed. Even without duplication at precisely the same dose, a low AD uncertainty will be estimated by jackknifing if there is more than one high leverage point with a high precision, no matter how inappropriate the model is. The simple delete-one jackknife itself cannot be used to assess the validity of a model, as even a very poor model will give a small uncertainty in AD provided the experimental uncertainty on each point is low. (A test for model aptitude has been given in a previous section.)

Cautionary measures

The following procedures should help to avoid some of the serious pitfalls in curve fitting and error estimation outlined above.

(i) A sufficient number of data points should be determined to allow for unusual points to be identified and the response curve to be well-defined. It would be unlikely that reliable results would be obtained with fewer than 8-10 observations (3-4 per parameter fitted), although this is very much dependent on the complexity of the model, the quality of the data and the method of analysis employed.

(ii) Leverage should be taken into account by concentrating data in the lower dose range, with some points at higher doses to give an indication of how the dose response curve behaves at higher levels. This concentration is particularly effective if the data has a strong linear component at lower doses but departs from linearity at higher doses, as for example in the linear/exponential model proposed by Levy (1989). Where the model is well defined, precise high dose data is extremely valuable, but where the model is less reliably known, even very precise high dose data should be downweighted to counteract the effects of leverage (Fox et al., 1980).

(iii) Allowance should be made for unequal variance in the Y-values, if it is considered significant. For example, the data can be weighted by the reciprocal of the variance for each data point, if it is known. Transforming the data by taking the log of the dependent variable and fitting a linear regression to the transformed data is mathematically equivalent to weighting an exponential data set for a constant percent error (de Levie, 1986).

(iv) The natural signal intensity should be evaluated as reliably as possible and weighted heavily in the curve fitting, as this point is fundamental to the concept of ESR and TL dating.

(v) The data should be checked for general form. Only if the data are linear (see Franklin, 1986) and natural signal intensity sufficiently well defined, can constrained weighted linear regression be used. The importance of the choice of model is reinforced by the example given in Grün and MacDonald (1989), who caution that apparently linear data may give very different results when fitted with an exponentially saturating curve, although the difference is probably not so severe if a combined linear/exponential curve is fitted.

(vi) If the data are of sufficiently high quality and quantity (>15 points, <5% uncertainty), quasi-likelihood methods as described by Berger *et al.* (1987), may be used to fit either linear or non-linear models.

(vii) If the data are non-linear and its quality/quantity deficient for likelihood techniques, an estimate of the uncertainty in AD can be obtained by the jackknife technique. Its interpretation should include a careful consideration of the potential problems of influence and leverage, and appropriate weighting to minimize their effects, particularly if the data have unequal variances. If jackknifing is used to estimate the uncertainty in AD, averaged values should be used for replicated points, and the variance in the mean value be included in the weighting. The jackknifed AD_i estimates should be checked for outliers, their significance assessed and any implications for calculating the best estimate considered. AD itself should be derived from the best fit to all data points and not by averaging the jackknife estimates of parameters.

(viii) In all curve fitting, particularly jackknifing, the data should be plotted along with their fitted curve to check for oddities, bearing in mind the possible effects of influence. Whether a significantly better fit has been achieved by using a model with more parameters can be checked by using the tests outlined above in the section on assessing model aptitude. It should be noted that even if a more complex model fits the curve more closely, extrapolation may still not be valid (e.g. if polynomials are fitted): for extrapolation, in contrast to interpolation, it is essential that the curve form expresses the underlying physical properties.

It is important to be pragmatic about the power of the techniques chosen: there are times when the quality of data may not justify the use of sophisticated techniques whose basic assumptions may not be satisfied, and less demanding techniques, even visual fitting, may be appropriate. The major source of systematic error is still likely to be the use of an incorrect model, and this cannot be detected by routine uncertainty calculations. Thus, at this stage, major progress in improving the reliability and accuracy of ESR and TL ages estimates is most likely to arise from theoretical advances in our knowledge of the appropriate curve forms to use.

Acknowledgements

We are grateful to Dr Rainer Grün for making the FITT progam available and initiating the use of re-sampling techniques to address a difficult problem in ESR and TL dating. We would like to thank our colleagues in the Department of Statistics, University of Auckland, particularly Prof. George Seber, Prof. Alistair Scott and Dr Alan Lee, as well as Dr Robert Gentleman, from McMaster University, for helpful discussions on the statistical principles involved. Thanks are also due to Prof. John Prescott for his meticulous refereeing; his comments have substantially clarified the arguments in this paper.

References

Apers, D., Debuyst, R., DeCanniere, P., Dejehet, F. and Lombard, E. 1981 A criticism of the dating by electron paramagnetic resonance (ESR) of the stalagmitic floors of the Caune de l'Aragoat Tautavel. In *Absolute Dating and Isotope Analysis in Prehistory- Methods and Limits* (eds DeLumley, H. and Labeyrie, J.) *Proc Pretirage*, 533-550

Berger, D.J. and Huntley, D.J. (1986) Linear regression of TL data. *Ancient TL* 4, 2, 26-29

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

Berger, D.J. (1990) Regression and error analysis for a saturating-exponential-plus linear model. *Ancient TL* **8**, 3, 23-25

Cook, R.D. and Weisberg, S. (1982) *Residuals and influence in regression*. Chapman and Hall, NewYork

Carroll R.J. and Ruppert, D. (1986) Discussion on 'Jackknife, bootstrap and other resampling methods'. *The Annals of Statistics* **14**, 4, 1298-1300

De Levie, R. (1986) When, why, and how to use weighted least squares. *J. of Chemical Education* **63**, 1, 10-13

Efron, D. (1982) The jackknife, bootstrap and other resampling plans. CBMS-NSF Reg. Conf. Ser in *Appl. Math. No 38. SIAM*, Philadelphia

Fox, T., Hinkley, D., and Larntz, K. (1980) Jackknifing in nonlinear regression. *Technometrics*, **22** 29-33.

Franklin, A.D. (1986) Extrapolation errors in linear regression. *Ancient TL* 4, 2, 31-35.

Freedman, D.A. (1986) Discussion on 'Jackknife, bootstrap and other resampling methods', *The Annals of Statistics* **14**, 4, 1305-1307.

Grün, R. and MacDonald, P.D.M. (1989) TL/ESR dose response curves. *Applied Radiation and Isotopes* **40**, 10-12, 1077-1080.

Grün, R. and Rhodes, E.J. (1991) On the selection of dose points for saturating exponential ESR/TL dose response curves *Ancient TL* **9**, 3, 40-46

Hall, P. (1986) Discussion on 'Jackknife, bootstrap and other resampling methods', *The Annals of Statistics* **14**, 4, 1311.

Hinkley, D. V. (1977) Jackknifing in unbalanced situations. *Technometrics* **19**, 285-292.

Miller, R.G. (1974) The jackknife- a review. *Biometrika* **61**, 1-15.

Misaqi, F.L. (1975) Applications of statistics to radiation surveys in mines. *MESA Information Report 1020*, US Dept of Interior, Washington.

Levy. P.W. (1989) Principles determing the length of time materials can be dated by TL, ESR and other trapped charge build-up methods. In *Occasional Publication no 9*, The Res. Lab. for Archaeology and the History of Art, Oxford University, pp24-29.

Lyons, R.G. (1988) A simple statistical method for comparison of accumulated dose in ESR dating. *Nuclear Tracks and Radiation Measurements* **14**,1-2, 361-363

Prescott, J.R. (1983) Thermoluminescence dating of sand dunes at Roonka, South Australia. *PACT* **9**, 505-512.

Rao, J.N.K. and Prasad, N.G.N. (1986) Discussion on 'Jackknife, bootstrap and other resampling methods', *The Annals of Statistics* **14**, 4, 1320-1321.

Readhead, M.L. (1984) *Thermoluminescence dating of some Australian sedimentary deposits* Unpub PhD thesis Australian National University.

Readhead, M.L. (1988) Thermoluminescence dating study of quartz in aeolian sediments from South-Eastern Australia *Quaternary Science Reviews* **7**, 257-264.

Rendell, H.M. (1985) Problems with linear regression as applied to TL data. *Ancient TL* **3**, 6-9.

Shao, J. (1986) Discussion on 'Jackknife, bootstrap and other resampling methods', *The Annals of Statistics* **14**, 4, 1322-1325.

Scott, E.M. and Sanderson, D.C.W. (1988) Statistics and the additive dose method in TL dating. *Nuclear Tracks and Radiation Measurements* **14**,1-2, 345-354.

Seber, G.A.F. (1977) *Linear regression analysis*. Wiley and Sons, New York.

Seber, G.A.F. and Wild, C.J. (1989) *Nonlinear regression*. John Wiley and Sons, New York.

Siminoff, J.S. and Chih-Ling Tsai. (1986) Discussion on 'Jackknife, bootstrap and other resampling methods', *The Annals of Statistics* **14**, 4, 1326-1328.

Wu, C.F.J. (1986) Jackknife, bootstrap and other resampling methods *The Annals of Statistics* **14**, 4,1261-1294.

PR. J.R. Prescott

Simulations of saturating exponential ESR/TL dose response curves - weighting of intensity values by inverse variance

Rainer Grün† and Edward J. Rhodes§
†Radiocarbon Dating Research Unit, RSPacS, Australian National University, G.P.O. Box 4, Canberra ACT 2601, Australia.
§Dept. of Geography, Royal Holloway and Bedford New College, Egham Hill, Egham TW20 0EX, UK.

Introduction

This paper presents the continuation of numerical simulations that were presented by Grün and Rhodes (1991). These simulations were used to investigate the influence of various parameters (even/exponential dose distribution; maximum irradiation dose, D_{max}, precision of TL intensity, and number of data points) on the determination of the equivalent dose, D_E, and the associated error assuming equal weights for the ESR/TL intensity. It was concluded that the size of the D_E error is critically dependent on the precision of the measured ESR/TL intensity and also on the selection of the maximum irradiation dose and the dose step distribution. It was commented by Berger (1991) that the assumption of equal weights for the ESR/TL intensity may invalidate the derived conclusions. He suggested further tests for unequal weighting, specifically weighting by inverse variance. These simulations have been carried out and are presented in this contribution.

The technical and experimental procedures of this paper are identical to Grün & Rhodes (1991) to which the reader is referred. Parameters are defined as follows:

D_E	equivalent dose
D_0	characteristic saturation dose
D_{max}	maximum radiation dose
I	measured ESR/TL Intensity
I_{max}	maximum intensity when all traps are filled
s.d.	standard deviation of computed D_E distribution (random error)
dev.	deviation of mean value of D_E distribution from preset D_E (systematic error)
s.d.(I)	uncertainty of ESR/TL intensity

Two dose distribution models were investigated:

Model 1:	even dose spacing, (e.g. 0, 2, 4, 6, 8....)
Model 2:	doubling of an initial dose step (e.g. 0, 1, 2, 4, 8....).

Simulations were performed for various preset D_E/D_o ratios of 0.003, 0.015, 0.03, 0.15, 0.3 and 1.5 (for D_0 = 3333 Gy, these values correspond to dose values of 10, 50 100, 500, 1000 and 5000 Gy, respectively). The results in this paper are based on 2000 randomly generated sets of ESR/TL values.

Results

Effect of the maximum irradiation dose, D_{max}

As in the previous paper, this experiment is based on ten ESR/TL data points per dose response curve including the natural sample with an assumed uncertainty in the ESR/TL intensity of 2%.

Figure 1 shows the standard deviation (s.d.) in the determination of D_E and the systematic deviation of the computed mean D_E value from the preset D_E. In all cases where $D_E < D_0$ (fig. 1A-E, left), the error in D_E estimation decreases rapidly with increasing D_{max} values, reaching relatively low levels when $D_{max} \geq 10D_E$. Above a D_E/D_0 ratio of 0.15 (fig. 1D-F, left), the error in D_E becomes more critically dependent on the choice of D_{max}, because the plot of s.d.(D_E) versus D_{max}/D_E shows clear minima. In case $D_E > D_0$ (fig. 1F, left), the error in D_E determination is at least 20% for $D_{max} \approx 2\ D_E$.

In the range $D_E<0.03D_0$ systematic underestimations occur when $D_{max}<10D_E$ (fig. 1A-C, right), whereas in the D_E range from 0.15 to $0.3D_0$ systematic overestimations occur when using $D_{max}>10D_E$ (fig. 1D-F, right). A selection of $D_{max} \approx 10D_E$ leads to the smallest overall error (s.d. in D_E determination plus systematic error).

Effect of dose distribution model

The exponential dose distribution (model 2) resulted in somewhat smaller errors (5 to 15%) in D_E determination than the even dose distribution (model 1).

Effect of the precision of the ESR/TL intensity

Figure 2 shows the errors in D_E estimation for different uncertainties in the ESR/TL measurement (D_{max} = $10D_E$, except for $D_E = 1.5D_0$, here $D_{max} = 2D_E$). The uncertainty in D_E determination is strongly dependent on the intrinsic uncertainty of a ESR/TL measurement. A four-fold improvement in the precision reduces the error in D_E by a factor of about 3.5. D_E values close to saturation can be measured with a reasonable degree of confidence only if the precision of the ESR/TL measurement is better than 1%.

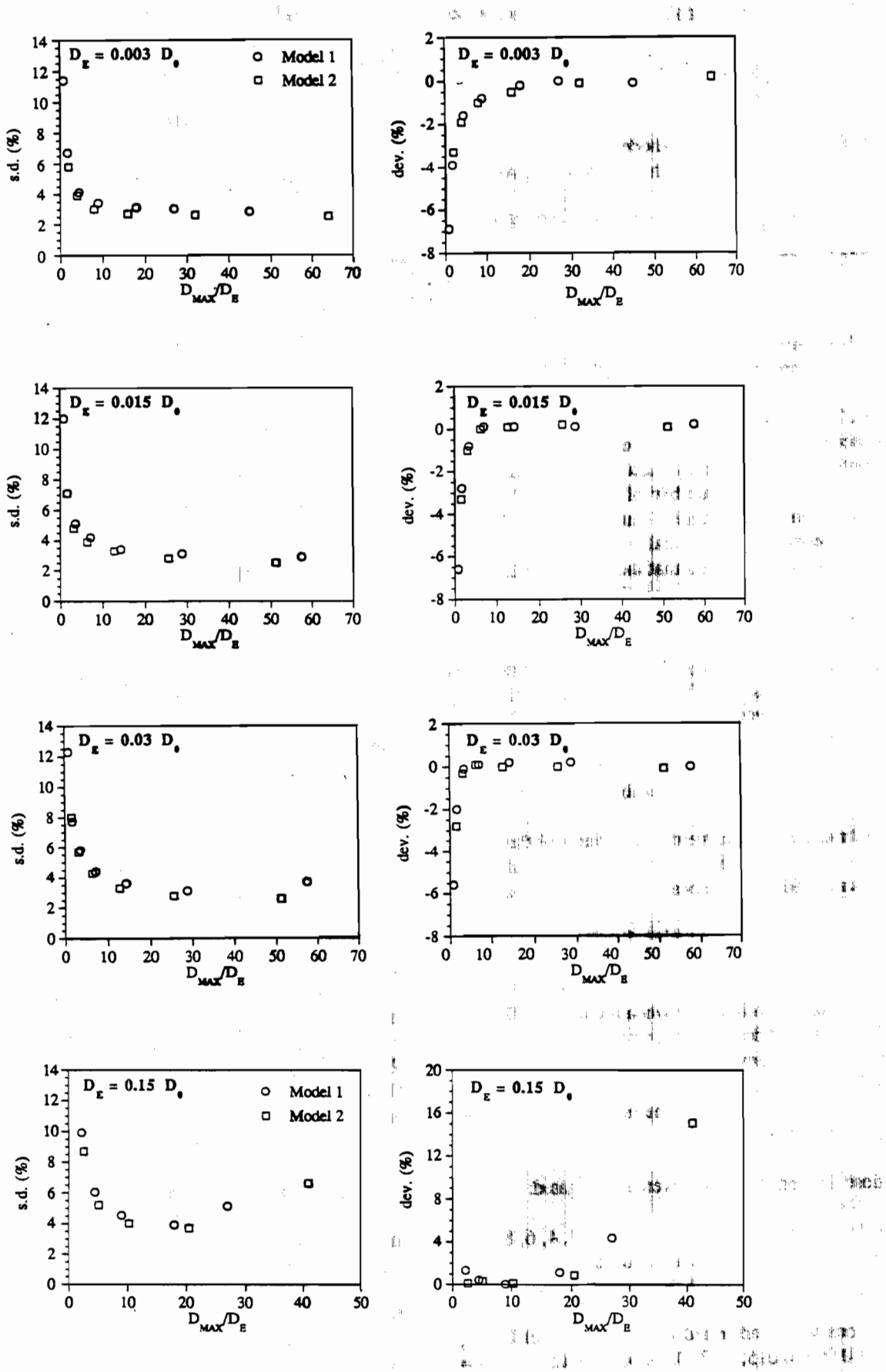


Figure 1. (A-D from top).
Standard deviation, s.d., (left) and systematic deviations (right) in D_E determination for preset D_E values of 0.003, 0.015, 0.03, 0.15, 0.3 and 1.5 D_0, 10 data points (one per dose) and a precision in ESR/TL intensity of 2%. Equal dose spacing (circles) and exponential dose spacing (squares).

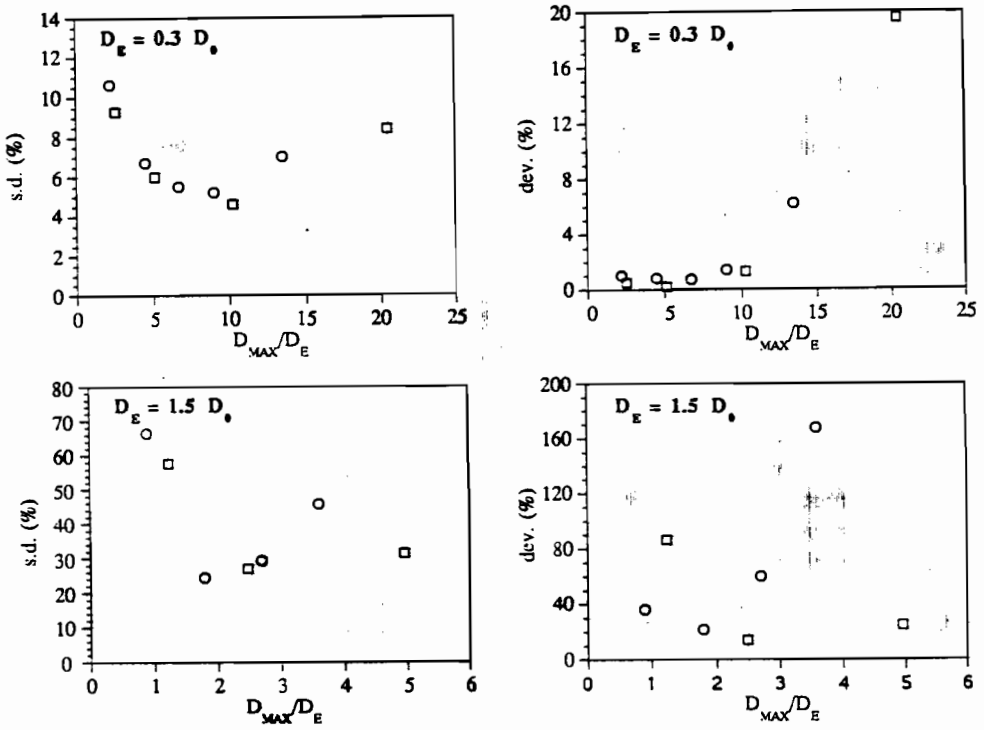


Figure 1. (E-F from top).
Standard deviation, s.d., (left) and systematic deviations (right) in D_E determination for preset D_E values of 0.003, 0.015, 0.03, 0.15, 0.3 and 1.5 D_0, 10 data points (one per dose) and a precision in ESR/TL intensity of 2%. Equal dose spacing (circles) and exponential dose spacing (squares).

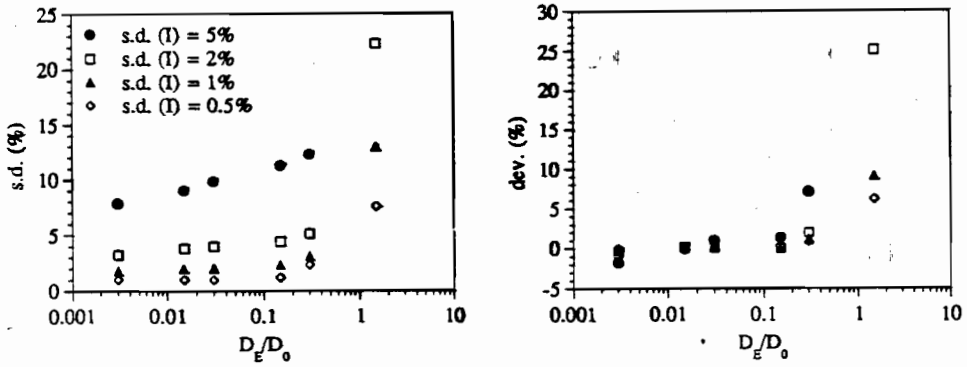


Figure 2.
Standard deviation, s.d., (left) and systematic deviations (right) in D_E determination for different precision in ESR/TL intensity, 10 data points, model 1 and $D_{max} = 10D_E$, except for $D_E = 1.5D_0$: $D_{max} = 2D_E$.

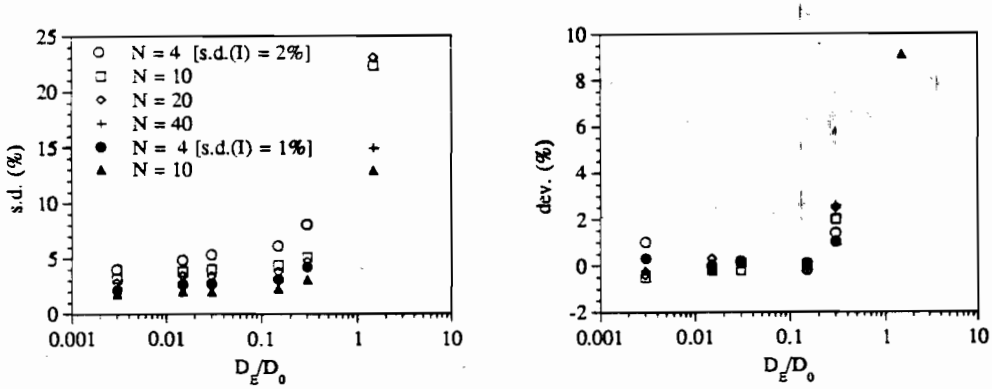


Figure 3.
Standard deviation, s.d., (left) and systematic deviations (right) in D_E determination for different number of data points, model 1 and $D_{max} = 10D_E$, except for $D_E = 1.5D_0$: $D_{max} = 2D_E$.

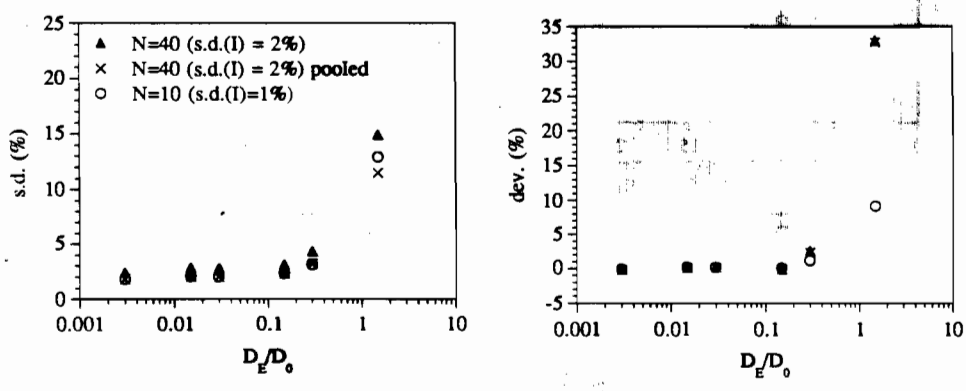


Figure 4.
Standard deviation, s.d., (left) and systematic deviations (right) in D_E determination for 40 data points, equally spaced and pooled at 10 equally spaced data points, respectively (2% precision), compared to 10 equally spaced data points with 1% precision. $D_{max} = 10D_E$, except for $D_E = 1.5D_0$: $D_{max} = 2D_E$.

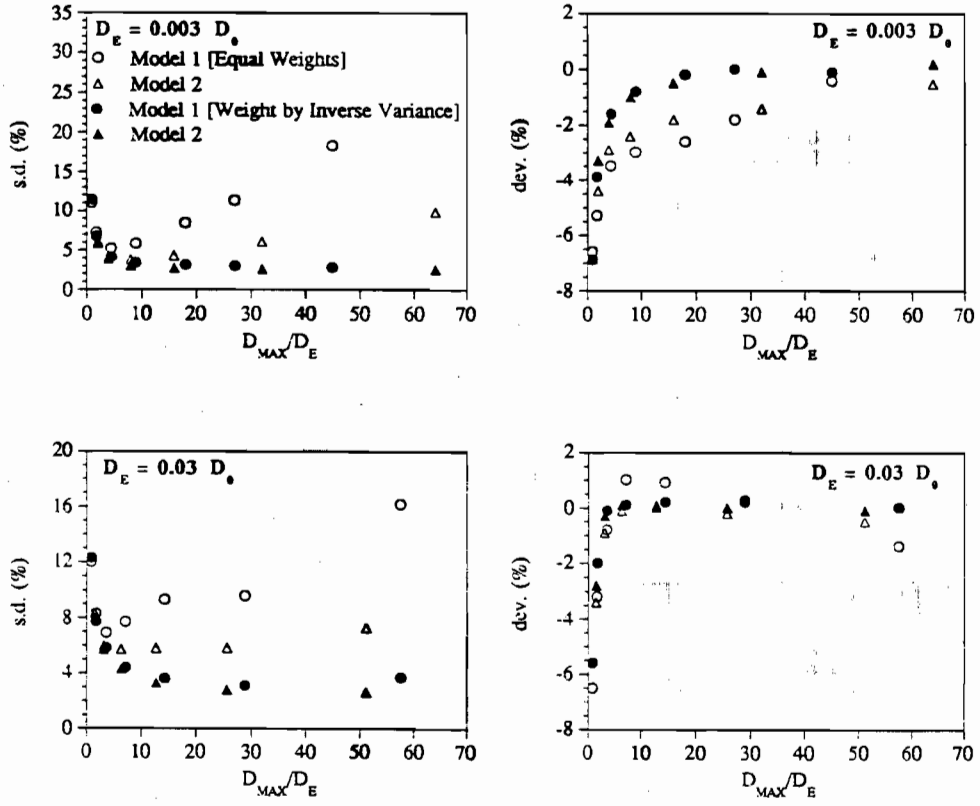


Figure 5.
Standard deviation, s.d., (left) and systematic deviations (right) in D_E determination for preset D_E values of 0.003 and $0.03D_0$ for 10 data points and a precision in ESR/TL intensity of 2%. Equal weights: open symbols, weight by inverse variance: closed symbols. Equal dose spacing (circles) and exponential dose spacing (triangles).

Effect of number of data points

Figure 3 shows the effect of the number of data points at equally spaced dose points within a given dose range (D_{max} = $10D_E$, except for D_E = $1.5D_0$, here D_{max}= $2D_E$). For most of the D_E range (0.003 to 0.3 D_0), a tenfold increase of the number of data points (from 4 to 40) reduces the D_E error by about 70%; a four-fold increase from 10 to 40 data points decreases the error by about 30%.The errors in D_E estimation using 4 data points with 1% uncertainty are as small as the errors obtained from 40 data points with 2% uncertainty. This result is very similar to the results that were obtained for equal weighting.

Pooling

Berger (1991) suggested that pooling of ESR/TL data points (repeated measurements at one dose step) would not be advantageous over measuring the same number of data points at different doses. When the values are pooled, the results are closely similar to those computed for higher precision of the ESR/TL intensity: i.e. if forty data points with an intrinsic uncertainty of 2%. each are equally distributed on 10 different dose steps (D_{max} = $10D_E$), the errors in D_E are nearly the same as for the simulation with 10 points and an intrinsic uncertainty of 1% (in the D_E range of 0.003 to 0.3 D_0; see figure 4). The systematic deviation is negligible in the D_E range of 0.003 to 0.3 D_0. The pooling procedure causes an improvement of the D_E error by about 40% when compared to the same number of measurements at different dose points. These simulation results do not substantiate Berger's comments in this particular respect.

Effect of weighting model.

The differences between the weighting models are shown in figure 5. The error in D_E determination using equal weighting shows distinct minima, especially for model 1 as shown for D_E = 0.003 and 0.03 D_0, respectively. The errors in D_E using weighting by the inverse variance are (i) far less dependent on the selection of the maximum irradiation dose and (ii) are smaller than the errors using equal weights. The systematic errors for weighting by the inverse variance are significantly smaller at D_E = $0.003D_0$ and are less dependent on the choice of D_{max}.

Discussion

The results of the simulations in this paper show some significant differences from the previous ones using equal weights. When using weights by inverse variance, the overall errors in D_E estimation are smaller. However, this is hardly surprising, because unfortunately the previous simulations contained a basic conceptual flaw: relative precisions were used for generating intensity values (i.e. the weights of the mean values were inversely proportional to the variance), but equal weights were assigned to the intensity values in the regression procedure. This means that quasi deliberately the wrong weighting model was used for the fitting of the data sets. Revised test runs have been carried out using intensity values generated in a fixed absolute range (i.e. the generated mean values have equal weights). Here, equal weight extrapolations lead to smaller overall errors than weighting by inverse variance. In other words, if the uncertainties in the ESR/TL intensity measurement are dominated by random errors, causing constant relative uncertainties, weighting by inverse variance seems the appropriate model for the fitting procedure and will lead to smaller errors in D_E determination than equal weighting. However, if the uncertainties in the TL/ESR intensity measurement are governed by systematic errors (background, constant interferences), causing constant absolute uncertainties, equal weighting seems the proper model and will lead to the correct D_E estimation. This clearly demonstrates the model-sensitivity of the extrapolation procedures. It has to be the aim of systematic studies to find the weighting model that describes the situation in ESR and TL most closely. Most uncertainties in ESR measurements seem to be random (weight, packing, orientation etc.), others maybe systematic (background, interferences by non-radiation sensitive signals).The situation in TL seems more complicated as discussed by Berger & Huntley (1989).

Conclusion

This study shows that the error in D_E estimation is dependent on use of the correct weighting procedure. Systematic studies have to be carried out to verify the intrinsic uncertainties of ESR/TL measurements.

This study supports the conclusions drawn in our previous paper (Grün & Rhodes 1991), that (i) the size of the error in D_E determination is critically dependent on the precision of the measured ESR/TL intensity; (ii) the maximum irradiation dose in the range of $10D_E$ leads to the smallest overall errors; (iii) exponential dose step distributions lead to smaller errors than equal dose distributions; and (iv) if additional measurements are carried out to improve the error in D_E estimation, it is advantageous to pool these, rather than measuring at additional dose steps in the same range. We want to emphasize again that the results are only valid for single saturating exponential dose response curves.

Acknowledgements

RG wishes to thank Mrs. J. Papps and Mrs. F.M. Grün for corrections on the manuscript. We thank the referee, GWB, for many helpful comments.

References

Berger, G.W. (1991) Reviewer's comments on Grün, R. and Rhodes, E.J. On the selection of dose points for saturating exponential ESR/TL dose response curves. *Ancient TL* **9**, 46.

Berger, G.W. and Huntley, D.J. (1989) Test data for exponential fits. *Ancient TL* **7**: 43-46.

Grün, R. and Rhodes, E.J. (1991) On the selection of dose points for saturating exponential ESR/TL dose response curves. *Ancient TL* **9**, 40-46.

PR Reviewer's Comments (G.W. Berger)

This is a commendable contribution to our further understanding of the effects of the several variables affecting the precision and accuracy of D_E estimates obtained from regressions to a saturating exponential model of TL/ESR dose-response curves. It is reassuring that some of the empirical procedures (e.g., dose doubling in growth-curve construction) long used by some of us now have a quantifiable justification. It is equally useful that the authors draw our attention again to the importance of weighting schemes. Their observation that different weighting schemes are appropriate for different error types (systematic or random) is significant. Berger and Huntley (1989) also distinguished the effects of these two error classes in their qualitative discussion on D_E plateaus, but Grün and Rhodes extend this distinction to its effects on modelling of growth curves.

A final comment concerns the fourth conclusion of the authors. The model (weighted saturating exponential) and procedures of Berger et al. (1987) (BLK87) do not suggest that pooling could be as significant as these numerical simulations suggest. Two practical examples of the effects of pooling with the BLK87 model are given here (fig.1., table 1), to demonstrate that there appear to be significant differences between their approach and that of Grün and Rhodes.

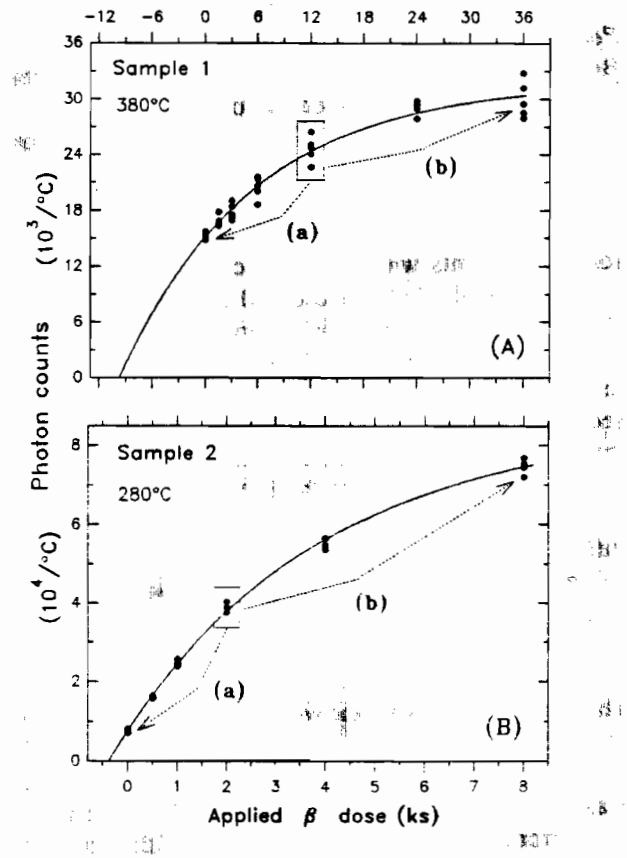


Table 1. Effect of pooling (fig.1)

	Parms.	Hole*	(a)**	(b)**	Normal**	(c)ǂ	(d)
SAMPLE 1	D_E††	10060	10160	9867	9698	9382	9196
	σ	1190	1100	1160	990	833	1045
	Var.‡	4.45%	4.55%	4.58%	4.57%	4.04%	4.51%
SAMPLE 2	D_E	388.5	388.0	388.2	377.6	378.9	377.7
	σ	10.1	11.7	11.0	10.9	9.53	10.2
	Var.	3.12%	3.25%	3.12%	3.45%	3.49%	3.31%

Notes.

* Boxed dose points deleted

** 'a', 'b' moved (and scaled) as in fig. 1, 'normal' means boxed points included.

ǂ Zero-dose points cloned (6 for each sample).

† Maximum-dose points cloned (5 for sample 1, 4 for sample 2).

†† D_E and σ in seconds.

‡ Scatter of data, from equation 4 of Berger et al. (1987)

These examples are representative of both large (fig. 1a) and small (fig. 1b) relative extrapolations. In fig. 1a 5-6 data points occur at each dose point; in fig. 1b, only 4 (except 6 at zero dose). The solid curves represent regression through all data points (the 'normal' situation in table 1).

As a first-order test of the effects of pooling, we may delete one set of data points (boxed in each fig.) and ask: 'Where should one place any additional points?'. This could simulate the situation faced in conclusion (iv) of Grün and Rhodes. Table 1 gives the results of this simple test. If the 'additional' points are pooled (and scaled) at either the low-dose (case 'a') or high-dose (case 'b') ends of the growth curves, there is no significant reduction in error in D_E compared to the absence of these data points (see σ values for first three columns, Table 1). There is, however, some reduction in error if we place these points at an additional dose step within the chosen dose range (e.g., within the boxes) (see 'normal' column), but only for sample 1. Furthermore, additional pooling beyond this appears to be (slightly) beneficial only if the pooling occurs at the low-dose points (compare columns 'c' and 'd'). This is perhaps not surprising, given the use of an inverse-variance weighting scheme.

This simple practical test illustrates that model differences are important, but does not detract from the welcome insights provided by the numerical simulations of the authors.

Berger, G.W., and Huntley, D.J. (1989) Treatment of error in plateau values - caveat emptor. *Ancient TL* 7, 27-29.

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

Authors' Reply

The two models in our paper that Berger refers to are: (i) all intensity measurements at different doses (ii) repeated TL/ESR measurements at a much reduced number of different doses. We do not expect to see any quantifiable effects if one moves only a very small number of measurements to other doses.

Bibliography

Aitken, M.J. and Valladas, H. (1992) Luminescence dating relevant to human origins. *Phil. Trans. R. Soc. Lond. B* **337**, 139-144.

Anderson, M., Jeannin, M., Rendell, H.M., Tardot, A. and Townsend, P.D. (1990) TL spectra of mineral mixtures: discrimination between different components. *Nuclear Tracks and Radiation Measurements* **17**, 569-577.

Calderon, T., Khanlary, M.R., Rendell, H.M. and Townsend, P.D. (1992) Luminescence from natural fluorite crystals. *Nuclear Tracks and Radiation Measurements* **20**, 475-485.

Hutton, J.T. and Prescott, J.R. (1992) Field and laboratory measurements of low-level thorium, uranium and potassium. *Nuclear Tracks and Radiation Measurements* **20**, 367-370.

Lees, B.G., Lu, Y.C. and Price, D.M. (1992) Thermoluminescence dating of dunes at Cape St. Lambert, East Kimberleys, northwestern Australia. *Marine Geology* **106**, 131-139.

Luff, B.J. and Townsend, P.D. (1992) High sensitivity thermoluminescence spectrometer. *Measurement Science Technology* **3**, 1-7.

Mejdahl, V. (1991) A survey of archaeological and geological samples dated in 1990. Geoskrifter nr 40, Aarhus University, 35 pp.

Parks, D.A. and Rendell, H.M. (1992) Thermoluminescence dating and geochemistry of loessic deposits from south-west England. *Journal of Quaternary Science* **7**, 99-107.

Schulmeister, J. and Lees, B.G. (1992) Morphology and chronostratigraphy of a coastal dunefield; Groote Eylandt, northern Australia. *Geomorphology* **5**, 521-534.

TL dates in:

Bryant, E.A., Young, R.W., Price, D.M. and Short, S.A. (1992) Evidence for Pleistocene and Holocene raised marine deposits, Sandon Point, New South Wales. *Australian Journal of Earth Sciences* **39**, 481-493.

Lees, B.G. (1992) Geomorphological evidence for late Holocene climatic change in northern Australia. *Australian Geographer* **23**, 1-10.

Mangerud, J. and Svendsen, J.I. (1992) The last interglacial/glacial period on Spitsbergen, Svalbard, *Quaternary Science Reviews* **11**, 633-664.

Nanson, G.C., Price, D.M. and Short, S.A. (1992) Wetting and drying of Australia over the past 300 ka. *Geology* **20**, 791-794.

Compiled by Ann Wintle

Letters

•*Vlasov & Kulikov's method*

In their paper published in 1989, Vlasov and Kulikov (*Physics and Chemistry of Minerals* **16**, 551-558, 1989) provide the most comprehensive description yet of the method they use to obtain TL dates from quartz. Since some of these are over a million years this is obviously of great interest. The paper is not particularly accessible and I thought it would be worthwhile to provide a summary to the best of my ability. At the end I shall make some comments of my own. The reader is advised that I may well have misinterpreted some aspects of the paper and that what follows should not be regarded as a faithful representation of the original, but is my interpretation.

The following refers to the light sum (i.e. TL intensity) for the "high-temperature" quartz peak at 310 °C. They show that it is necessary to anneal their irradiated samples at 200 °C for 20 minutes in order to separate this peak.

(a) The rate equation used is:

$$dn/dt = (n_o - n)^2 . \sigma_1 . P_{eff} - n/\tau$$

Here the first term represents trap filling by the radiation and the second term represents thermal trap emptying.

n	is the concentration of filled traps at time t,
n_o	is the total concentration of traps,
P_{eff}	is the dose-rate,
σ_1	is the trapping cross section, and
τ	is the mean trapped charge lifetime due to thermal emptying.

(b) The equation is solved for the case of laboratory environmental irradiation in which case the second term is omitted. The solution yields the following relation between lab dose and light sum:

$$1/\Delta S = a + b/D$$

where, D = lab dose; $\Delta S = S_\infty - S$; S = light sum and S_∞ is its value in the limit $D = \infty$; $a = 1/(S_\infty - S_{nat})$; $b = a/(n_o - n_{nat})\sigma_1$

(c) Experiments of TL intensity (S) vs lab dose (D), i.e. N + γ data, are used to evaluate the parameters a and b using a plot of $1/\Delta S$ vs $1/D$ on which the data are linear. This permits determination of $n_o\sigma_1$. The range of gamma doses used is 500 - 10,000 Gy.

(d) The full rate equation is solved for the case of the natural dose rate and independently determined τ. A graph of S vs time is produced and two points placed on the curve. The first is the natural intensity (S_n), the second is the residual (S_o). The age is taken to be the difference between these on the time axis.

(e) The value of S_o is taken to be a particular fraction (f) of the value of S that would occur at $t = \infty$ under natural conditions. A table of f values for different sediments is given.

Comments

The glow curves and growth curves shown do not look like those for quartz that I am familiar with. The authors make a point of stating that HF is not used because it would alter the radiation sensitivity; perhaps there is a connection.

The form of the first term in the rate equation is most unconventional. It would be helpful if the authors would expand on the physical basis they state for it. The data shown agree extremely well with the linear form of the solution given above.

The definition given for n_o is the 'initial concentration of vacant traps'. This has the effect of making the rate equation itself dependent on the initial conditions; the definition I give above avoids the difficulty.

Reading the paper is made easier once one realizes that in equation 17 the symbol n represents $(n_o - n_{nat})$, that the left hand side of equation 20 should be $1/\Delta S$, and that the ordinate of Fig. 4a is ΔS^{-1}.

D. J. Huntley
Department of Physics, Simon Fraser University

*

• *Symbols Proposals by Martin Aitken*

I very much welcome Aitken's initiative to introduce a standardised set of symbols for TL and optical dating. I hope that it was also the aim to include ESR dating where appropriate.

The choice of some of the characters and combinations may be a matter of personal preference, however, I should like to raise the following points:

1) It seems unfortunate that dose values are expressed with four different capital characters, R, E, Q and P. The international symbol for energy dose is D (see *Handbook of Physics and Chemistry* and *The Symbols Committee of the Royal Society*, 1975). The pal(a?)eodose, as defined by Aitken, is expressed as 'effective dose' (i.e. corrected for alpha efficiency), and is actually the dose value that is equivalent to the kind of radiation that has been used in its determination (i.e. gamma or beta rays). There seems no reason to avoid the international symbol and to denote specific procedures by subscripts. D_E could be the equivalent dose as used in the age equation, D can stand generally for applied dose, D_o for saturation dose, and D_I for the supralinear correction. The relationship between measured ESR intensity, I, and dose would be:

$$I = I_{max}\,(1 - \exp(-(D+D_E)/D_o))$$

or in TL/ optical dating, using the other symbols suggested by Aitken:

$$L = L_i\,(1 - \exp(-(D+D_E)/D_o)$$

2) The selection of the character of the dose rate seems more problematic. The international symbol is D, which is difficult to print (one may want to use the German notation D' instead). However, both notations, D and D', suggest an instantaneous dose rate. However, most Quaternary samples show radioactive disequilibria and the cited dose rates are usually average dose rates over the calculated age of the sample. In these cases a notation implying a derivative is not strictly correct. A concept of using a dose system (i.e. multiplying all the dose rates with the calculated age) instead of dose rates may overcome this problem. The problem of disequilibria and the citation of dose rates involving these has to be addressed.

It is evident that the introduction of a standardised set of symbols for TL, ESR and optical dating requires a rigorous discussion.

Rainer Grün
RSPacS, Canberra

Author's Response

I am glad to have some comment on my proposals and hope that this may set the ball rolling for more widespread discussion. I have also received informal critical comment from D.J. Huntley, who like Grün, strongly favours the use of *D* for those (whereas I had proposed the use of *D* for dose-*rate*). Since, as Grün points out, the international symbol for energy dose is D, there is indeed a case for using *D* for laboratory-applied dose. However I strongly favour retention of *P* for paleodose (optionally, palaeodose) both for simplicity (see below) and since, again as pointed out by Grün, it is an effective dose (at any rate when used in the age equation) rather than a true energy dose.

The reason for which I proposed the use *E*,*Q* and *P* rather than using suffixes was to reduce the need for suffixes and double suffixes. On some PC's these, whether single or double, are wasteful of time; more importantly they are confusing when spoken, increasing the comprehension difficulties of non-English speakers.

As regards the exponential equations D_s is preferable to D_o since the latter implies some connection with zero, and suffix *i* is preferable to suffix *max* since it is shorter. I agree that it would be desirable to use *I* rather than *L* so as to bring TL/OD and ESR into line, with the proviso that one should avoid the use of a font (such as Geneva) in which there is confusion with a small *I* ('ell').

Certainly there is a need to evolve a system that deals with the problem of a changing dose-rate, such as due to uranium-uptake in teeth or calcite as well as due to radioactive disequilibrium. However I think possibilities should be developed by individuals and tried out in published papers before discussion by the (Conference) Committee. In the meantime, and in any case for the many sites on which assumption of a constant dose-rate is acceptable, I suggest that *R* be used to denote average dose-rate.

Martin Aitken
University of Oxford