Absorbed dose, equivalent dose, measured dose rates, and implications for OSL age estimates: Introducing the Average Dose Model

Guerin, Guillaume; Christophe, C.; Philippe, A.; Murray, Andrew Sean; Thomsen, Kristina Jørkov; Tribolo, Chantal; Urbanova, Petra; Jain, Mayank; Guibert, Pierre; Mercier, Norbert

Published in:
Quaternary Geochronology

Link to article, DOI:

Publication date:
2017

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Absorbed dose, equivalent dose, measured dose rates, and implications for OSL age estimates:
introducing the Average Dose Model

G. Guérin1,*, C. Christophe1, A. Philippe2, A.S. Murray3, K.J. Thomsen4, C. Tribolo1, P. Urbanova1, M. Jain5, P. Guibert1, N. Mercier1, S. Kreutzer1, C. Lahaye4

1 UMR 5060 CNRS - Université Bordeaux Montaigne, IRAMAT-CRP2A, Maison de l’archéologie, Esplanade des Antilles, 33607 Pessac cedex, France

2 Jean Leray Laboratory of Mathematics (LMJL), UMR6629 CNRS - Université de Nantes, France.

3 Nordic Laboratory for Luminescence Dating, Department of Geoscience, Aarhus University, DTU Risø Campus, DK-4000 Roskilde, Denmark.

4 Center for Nuclear Technologies, Technical University of Denmark, DTU Risø Campus, DK-4000 Roskilde, Denmark.

Abstract

Luminescence ages are calculated by dividing an absorbed dose by the dose rate to which the natural dosimeter has been exposed. In practice, one measures an equivalent dose, \(D_e\); in the absence of an alpha dose contribution, this should be indistinguishable from the dose absorbed in nature. Here we first review the relationship between absorbed dose, equivalent dose and dose rate, and the measurements that lead to their estimation; we restate that, in contrast to recent suggestions, an equivalent dose is not a physically different quantity from a beta or gamma dose absorbed by quartz grains. Statistical analysis of OSL data is of great importance when dealing with single grain data, since such data commonly exhibit significant scatter. However, dose rate measurements provide an arithmetic mean of dose rates absorbed by individual grains; in this article, we propose a new model to estimate the average dose absorbed by the grains. We thus introduce a new model for OSL age estimates: the Average Dose Model (ADM). We argue that ADM ages should be more accurate than Central Age Model (CAM) based ages, and we provide experimental evidence supporting this expectation. We also argue that the use of the Finite Mixture Model should be avoided. Finally, we discuss the implications for multi-grain age estimates derived from well-bleached samples.

Keywords: OSL data analysis; Dose rate measurements; Central Age Model; Average Dose Model

Highlights:

- Dose rate estimates correspond to arithmetic means
- OSL age models should thus aim at arithmetic means of absorbed doses
- We introduce the Average Dose Model (ADM)
- ADM ages for known-age samples are more accurate than CAM ages
- We argue against the use of the FMM
1. Introduction

In luminescence dating methods, an age is obtained by dividing the dose absorbed by a dosimeter from ionising radiation (this dose is known variously as the palaeodose, burial dose or archaeological dose), by the time-averaged rate at which dose has been absorbed since the last signal resetting event:

\[ t = \frac{\Delta}{\bar{D}}, \]  

where \( t \) is the age (in ka), \( \Delta \) is the absorbed dose (in Gy) and \( \bar{D} \) is the dose rate (in Gy.ka\(^{-1} \)).

Because the palaeodose is not directly measurable, several luminescence signals whose intensities vary as a function of dose are measured as proxies; thermoluminescence (TL: Aitken, 1985), Optically Stimulated Luminescence (OSL: Huntley et al., 1985), Infra-Red Stimulated Luminescence (IRSL: Hütt et al., 1988) are the most frequently used. Here we discuss what is measured in the process of dating sediment with such signals, and based on these measurements, how we derive the quantities needed to calculate an age. Since both the numerator and denominator in the age equation are average estimates based on various measurements of physical quantities (absorbed dose or dose rate), one cannot discuss the statistical analysis of one term without the other.

We first discuss the concept of equivalent dose, \( D_e \), widely used to describe the measurements made to estimate the absorbed dose, and the differing definitions of \( D_e \) that exist in the literature. Based on how dose rates are determined, we then argue for a change in the methods used to calculate the palaeodose, characteristic of a sample age, which is determined from observed individual \( D_e \) estimates made on aliquots; for this purpose, we introduce a new model for statistical analysis of \( D_e \) distributions: the Average Dose Model. Tests of this model in comparison with the Central Age Model (CAM: Galbraith et al., 1999) support a significant improvement of OSL ages for well-bleached samples when using the Average Dose Model. Finally, consequences for the use of other age models, and in particular of the Finite Mixture Model (Roberts et al., 2000), as well as for multi-grain OSL age calculation, are discussed.

2. The concept of equivalent dose

The equivalent dose (\( D_e \), in Gy) was originally defined as the beta or gamma laboratory dose that results in the same signal intensity as the natural signal, i.e. the signal induced by the absorbed dose. Here, it should be noted that we assume there were no residual charges left in the dosimeter of interest at the time of zeroing; in other words, we focus only on well-bleached samples. Aitken (1985) stated that palaeodose (see also Huntley, 2001) is the sum of an equivalent dose (his ED or Q) and an intercept (see his Fig. 2.1, p. 19). It now appears that the intercept, \( I \), was largely an artefact of the additive dose protocol used by Aitken for illustration; in the SAR protocol (Murray and Wintle, 2000), and thus in most current OSL and IRSL studies, this intercept does not exist (the dose response curve passes through the origin, see, e.g., Banerjee et al., 2001). Thus, applying Aitken’s definition to regenerative protocols, absorbed dose is identical to equivalent dose (his ED, our \( D_e \)). As a result, the equivalent dose, if measured accurately, is indistinguishable from the absorbed dose (i.e. the palaeodose).

We note in passing that there is no alpha dose contribution to burial dose in multi- or single-grain dating when the grains have been etched to remove the outer alpha-irradiated layer. This is standard practice in all coarse-grain quartz dating; as a result, in the following we do not consider alpha
irradiations (except from a generally small internal dose rate contribution, following e.g. Vandenberghe et al., 2008).

However Galbraith (2015) considers, in the context of single grain $D_e$ measurements, that absorbed dose and $D_e$ can be very different, due to natural variability in OSL properties: “for two grains that have absorbed the same radiation dose, their equivalent doses, even when measured very accurately, will typically differ because of their differing OSL properties” (Galbraith, 2015; see also Galbraith et al., 2005). Whereas in the past luminescence measurements have been used to derive a best estimate of $D_e$, and scatter in such measurements around the (unknown) absorbed dose has been viewed as arising from uncertainty in measurement (both random and systematic), Galbraith is proposing that the accuracy of each individual measurement of $D_e$ from each aliquot (single- or multi-grain) is only limited by the quantifiable uncertainties (in this case, random). In his view, each grain/aliquot may have a genuinely different $D_e$, not because it has absorbed a different dose, but because it has different luminescence characteristics.

We consider Galbraith’s view that the equivalent dose from a grain can be accurate while still being different from the absorbed dose to be fundamentally different from previous definitions and usages, both explicitly stated and implied; in the context of coarse-grain dating, our interpretation of these earlier definitions is that the two quantities are identical. Galbraith (2015) seems to treat equivalent dose as a physical quantity in itself, arising because of the different OSL attributes of the sample, while in our view equivalent dose is the measured value of the absorbed dose or palaeodose. Any differences between equivalent dose and absorbed dose originate from measurement; the existence of such differences simply indicates that the equivalent dose has been measured inaccurately and/or imprecisely, not that it is a fundamentally different quantity.

[Insert Table 1 and Fig. 1 here]

Based on this definition of equivalent dose (defined as the beta or gamma laboratory dose that results in the same signal intensity as the natural signal), we can now review the evidence concerning analytical uncertainties on luminescence-based $D_e$ values published in the literature. These uncertainties, often estimated with the Analyst software (Duller, 2015), usually include contributions from counting statistics, curve fitting, and instrument reproducibility. The rate of absorption of dose by individual grains in nature is unknown; only a mass-averaged value is derived from dosimetry measurements (see section 3 below). As a result, at the single grain level one can only discuss an experimental comparison between equivalent dose and absorbed dose in the case of dose recovery experiments (Thomsen et al., 2005). Table 1 lists the samples analysed (these samples were already analysed in Guérin et al., 2015b; for details, the reader is referred to this publication) and the results of dose recovery experiments, including the intrinsic overdispersion (OD$_{int}$, following the terminology of Thomsen et al., 2005). In such experiments, the latent OSL signals from samples are reset in some manner, usually by exposure to sufficient light, before a known dose is given under controlled laboratory conditions. Experimentally observed distributions of $D_e$ are then compared to this known dose. These distributions usually display greater dispersion than that explained by the quantifiable analytical uncertainties, i.e. non-zero overdispersion (Table 1, Fig. 1; see also e.g. Galbraith et al., 2005; Thomsen et al., 2005, 2012). If the known dose was absorbed from gamma radiation, all grains can be experimentally arranged to have absorbed the same dose, and so any overdispersion must arise from the measurement process. In beta dose recovery experiments, where one is comparing beta $D_e$ values with a given grain-specific beta dose (which does not need to be well known in absolute terms, i.e. in Gy), grains are exposed to the same electron spectra, at the same location on the disc, resting in the same position on the sample disc, etc., both for the given dose and for the regeneration doses; as a result many or most of the potential sources of scatter are avoided. Then any overdispersion in the $D_e$...
distribution is presumed to arise from unrecognised sources of measurement uncertainty. This OD\textsubscript{int} is often non-trivial compared to quantifiable sources of uncertainty, and we must deduce that, in the literature, analytical uncertainties attributed to experimental D\textsubscript{e} values are systematically underestimated, often by a large amount. This observation argues against the use of any graphical representation based on analytical uncertainties, whether radial plots (Galbraith, 1988) or abanico plots (Dietze et al., 2015); since the true measurement uncertainties are unknown, such plots display false information and should be avoided. If they must be used, then presumably best estimates of uncertainty (i.e., including OD\textsubscript{int}) should be used (as for example in Reimann et al., 2012; Guérin et al., 2016). To illustrate this point, Fig. 2 shows two different radial plots for sample BR-2011-8 (Lahaye et al., 2015); in Fig. 2a only the analytical errors are taken into account, while in Fig. 2b the intrinsic overdispersion, which is by definition an additional measurement uncertainty, is added in quadrature to the individual analytical errors. If visual interpretation is to be used for the analysis of this single grain D\textsubscript{e} distribution, Fig. 2a is clearly misleading.

In contrast to beta dose recovery experiments, in gamma-dosed samples all grains have been arranged to absorb the same dose. Although the radiation differs between the given gamma dose and the regeneration beta doses, both are low LET radiations, and it can reasonably be assumed that their ionisation densities are similar. However, one cannot completely exclude grain to grain variations in the beta dose irradiation geometry, and the total dispersion presumably also includes contributions from uncertainties arising from grain manipulation, disc preparation, sample loading etc. (see for example the beta dose recovery experiment of Thomsen et al., 2005, where they beta irradiated grains in the reader, removed the grains from the single-grain disc, and then reloaded them, to mimic a first irradiation in a geometry different from that during regeneration). Even if the beta source has been properly calibrated for each grain position, the OD\textsubscript{int} must include the unquantified contributions from such manipulation, and this may explain, at least in part, the generally higher dispersion in gamma dose recovery distributions compared to those from beta dose recovery experiments (Fig. 1; see also Thomsen et al., 2005; 2012).

This discussion is not simply concerned with semantic pedantry. The definition of D\textsubscript{e} is of direct importance to dating because we compare absolute doses with absolute dose rates; the dose rates are derived from fundamental nuclear data, and represent energy deposition rates. Because of this, we are interested in the energy actually absorbed by grains. The relationship between this absorbed dose and the experimentally determined D\textsubscript{e} (and associated uncertainties) is thus of fundamental importance in age calculation. In particular, one must distinguish between the different sources of scatter in observed D\textsubscript{e} distributions in order to identify appropriate statistical analysis. We regard the intrinsic overdispersion parameter, OD\textsubscript{int}, as describing unrecognised measurement errors in D\textsubscript{e}; this definition now allows us to move on to discuss the relationship between the D\textsubscript{e} (as a measurement of absorbed dose) and dose rate.

### 3. Measurements of dose rates

It is well-known that the variability in dose rates to single grains may be a significant source of dispersion in the D\textsubscript{e} distributions of well-bleached samples (Olley et al., 1997; Nathan et al., 2003; Mayya et al., 2006; Guérin et al., 2015b). This source of dispersion was shown by Guérin et al. (2015b) to be large enough to explain all the extrinsic overdispersion – following the terminology of Thomsen et al. (2005) – in a natural sample (the ‘intercomparison sample’ of Murray et al., 2015).

With this in mind, let us consider a sediment sample including n identical quartz grains. In practice, Eq. (1) cannot directly be used for each of these grains because the dose rate to each quartz
grain is experimentally inaccessible. Therefore, assuming that all grains have the same age \( t \), we can write based on Eq. (1) that for all \( j \) (\( \forall j \))

\[
\forall j, \; t = \frac{\Delta_j}{D_j},
\]

(2)

where \( \Delta_j \) is the dose absorbed by grain \( j \) and \( D_j \) is the dose rate to which this grain has been exposed. For clarity, Table 2 lists all variables used in the following equations.

\[\text{[Insert Table 2 here]}\]

Different methods are used for dose rate determination. Among the most frequently used are gamma spectrometry (both high resolution, laboratory-based, and in situ using portable scintillators), artificial dosimeters (e.g., \( \text{Al}_2\text{O}_3 : \text{C} \) pellets), alpha and beta counting techniques, etc. (for a summary, see e.g. Aitken, 1985). In the more commonly used nuclide-specific techniques (e.g., neutron activation analysis, gamma spectrometry) the concentrations (or activities per unit mass) of various members of the U and Th decay chains, and of \( ^{40}\text{K} \) are measured directly. These concentrations are then used to calculate the infinite matrix dose rate using dose rate conversion factors. The latter provide the average energy emitted per disintegration, and are derived from tables of nuclear data (e.g., Guérin et al., 2011). In the case of U- and Th-series, the average energy emitted per unit time and mass is summed over all daughters from \( ^{238}\text{U} \), \( ^{235}\text{U} \), and \( ^{232}\text{Th} \), respectively. In other words, we estimate the total amount of energy emitted in the sample per unit time and unit mass. Using the infinite matrix assumption (Roesch and Attix, 1968; Aitken, 1985), this rate is equal to the rate of energy absorption per time and mass. For simplicity, let us assume that we have a sediment entirely made up of \( n \) identical grains of quartz, all having the same mass and age \( A \), and that the radioactive sources have negligible mass. Since dose rate is the total amount of energy absorbed per unit time and mass in the sample, and since energy is a cumulative quantity, if one grain receives more than the average dose rate, say a fraction \((1+x)\) of the average, then the remaining \((n-1)\) grains receive, on average, a fraction \((1-x/(n-1))\) of the average dose rate. If another grain receives a fraction \((1-x)\) of the average dose rate, all remaining \((n-2)\) grains must receive the average. This statement can be generalised: no matter the distribution of dose rates to individual grains, the invariant parameter is the amount of energy available for the grains, independent of how the radioactivity is distributed in the sample (see Guérin et al., 2012b; Guérin et al., 2015b).

Since nuclide-specific techniques in general involve the comparison of a signal intensity with that from a standard, this is likely to lead to multiplicative error properties; in the simplest cases (e.g. gamma spectrometry analysis of \( ^{40}\text{K} \)) the ratio of peak areas (unknown divided by standard) is multiplied by the known concentration in the standard. Nevertheless, despite the nature of these quantifiable uncertainties, we estimate the average radionuclide concentration and so the arithmetic mean dose rate absorbed by individual grains. To illustrate this, consider the simulated dose rate distributions to single grains presented in Guérin et al. (2015b). Depending on whether the activity is distributed heterogeneously (in 200 µm grains) or homogeneously (in a clay matrix) the geometric mean dose rate is very different. But in both cases the dose rate determined by standard dose rate measurements will be the same (the average dose rate) because it is derived from the mass averaged radionuclide concentration.

Alpha, beta or gamma integral counting methods induce an additional systematic source of uncertainty not present in spectrometric methods, because the relationship between count rate and
dose rate is, to a greater or lesser degree, dependent on the relative proportions of radionuclides (which is, of course, in general unknown – for discussion of these dependencies, see e.g., Aitken, 1985; Ankjærgaard and Murray, 2007). But for all these methods, the calibration is performed using standards for which the known quantities are the radioelement contents of K, U and Th and daughters. Thus, independent of whichever dose rate estimation method is used, the calculated dose rate corresponds to the average (arithmetic mean) dose rate to individual grains:

$$\bar{D} = \frac{\sum D_j}{n}. \quad (3)$$

Based on Eq. (3) we can now write:

$$\forall j, \Delta_j = tD_j = t \frac{\sum D_j}{n}, \quad (4)$$

Then,

$$t = \frac{\Delta}{\bar{D}}, \quad (5)$$

where the sample palaeodose \(\Delta\) is defined considering the \(\Delta_j\) as independent and identically distributed (i.i.d.) random variables, and through the strong law of large numbers, as

$$\sum \frac{\Delta_j}{n} \xrightarrow{a.s.} \Delta = \mathbb{E}(\Delta_j), \quad (6)$$

where \(\mathbb{E}\) denotes the expected value. Thus, for age determination the aim of any statistical modelling of \(D_e\) distributions should be the expected dose \(\Delta\) absorbed by the grains.

4. Consequences for age calculation

With this aim in mind, we can now discuss statistical analysis of single grain equivalent dose distributions. We first describe the CAM and then introduce the Average Dose Model.

4.1. The Central Age Model

{Insert Fig. 3 here}

We consider a distribution of \(n\) equivalent doses and associated relative analytical uncertainties \((D_{e,j}, \sigma_j)\) as discussed in section 2. Fig. 3 shows a hierarchical representation of the CAM, which aims at calculating a central equivalent dose \(D_{e,CAM}\) and the dispersion \(\sigma\) of individual "true" (but see section 2) equivalent doses around the central \(D_{e,CAM}\). In the CAM, the observed equivalent doses satisfy the following equations:

$$d_j = \delta_j + \varepsilon_j \quad (7)$$
where \(d_j = \log(D_{e,j})\), \(\delta_j\) is the logged “true” equivalent dose, \(\varepsilon_j\) is an error term such that \(\mathbb{E}(\varepsilon_j) = 0\) and \(\text{var}(\varepsilon_j) = \sigma_j^2\), and

\[\delta_j = d + \sigma \eta_j\]  
(8)

where \(d = \log(D_{e,\text{CAM}})\) is the logged central equivalent dose and \(\eta_j \sim \mathcal{N}(0, 1)\). In other words, it is assumed that the logged individual equivalent doses are distributed according to a Gaussian distribution around the logged central equivalent dose. The unknown parameters are \(D_{e,\text{CAM}}\) and \(\sigma\).

Eqs. (7) and (8) can be compacted as:

\[d_j = d + \varepsilon_j + \sigma \eta_j\]  
(9)

or

\[D_{e,j} = D_{e,\text{CAM}} e^{\varepsilon_j + \sigma \eta_j}.\]  
(10)

Assuming that \(\varepsilon_j \sim \mathcal{N}(0, \sigma^2_j)\), then \(\mathbb{E}(D_{e,j}) = D_{e,\text{CAM}} \mathbb{E}(e^{\varepsilon_j + \sigma \eta_j}) = D_{e,\text{CAM}} e^{\frac{1}{2}(\sigma^2 + \sigma^2_j)}\). This is equivalent to assuming that

\[D_{e,j} \sim \log \mathcal{N}(d, \sigma^2 + \sigma^2_j).\]  
(11)

Note that the observed \(D_{e,j}\) are not identically distributed. To estimate \(d\) and \(\sigma\), the log-likelihood function (for example defined in Pawitan, 2001) is used:

\[
\log L(D_{e,1}, \ldots, D_{e,n}, \sigma_1, \ldots, \sigma_n, d, \sigma) = \sum_j - \log (\sqrt{2\pi} \sqrt{\sigma^2 + \sigma^2_j}) - \frac{(d_j - d)^2}{2(\sigma^2 + \sigma^2_j)}.
\]  
(12)

Thus, the maximum likelihood estimators \(\hat{d}\) and \(\hat{\sigma}\) of \(d\) and \(\sigma\) verify the following equations in \(d\) and \(\sigma\):

\[
d = \frac{\sum_j d_j}{\sum_j 1} - \frac{\sum_j 1}{\sum_j (\sigma^2 + \sigma^2_j)}
\]  
(13)

and

\[
\sum_j \frac{1}{\sigma^2 + \sigma^2_j} = \sum_j \frac{(d_j - d)^2}{(\sigma^2 + \sigma^2_j)^2}.
\]  
(14)

Since the exponential function is regular, \(e^{\hat{d}}\) is the maximum likelihood estimator of \(e^d = D_{e,\text{CAM}}\), which is the median of the lognormal distribution from Eq. (11). It can be noted here that this parameter is also the geometric mean of the lognormal distribution if the \(D_{e,j}\) are identically distributed; indeed Eq. (13) corresponds to a weighted geometric mean of individual equivalent doses.

For simplicity, consider a case where the analytical uncertainties (\(\sigma\)) and the intrinsic overdispersion are negligible compared to the dispersion in dose rates to single grains, i.e. the
dispersion parameter in the CAM is the extrinsic overdispersion, and it is mainly due to dose rate heterogeneities. In such a case, Eq. (13) will give the unweighted geometric mean of individual equivalent doses \( (d) \). Since here we assume that measurement uncertainties (including the intrinsic overdispersion, cf. section 2) of equivalent doses are negligible, this is the same as the geometric mean of the doses absorbed by the grains. To illustrate this, Fig. 4 shows a radial plot for sample BR-2011-8 (the same sample as that used for the radial plots in Fig. 2) in which the uncertainties include the CAM OD. For this sample, the high overdispersion (60 \%) is the dominant dispersion factor in the \( D_e \) distribution. As a result, almost all individual \( D_e \) estimates are given the same weight in the calculation of the weighted geometric mean (Eq. 13), which in this case amounts to an unweighted geometric mean. Thus, because the geometric mean of a distribution is always less than or equal to its arithmetic mean, we may write that

\[
t \geq \frac{D_{e,CAM}}{D}.
\]  

This inequality can be generalised whenever the extrinsic overdispersion is not null (provided that the measurement uncertainties, including the intrinsic overdispersion, display multiplicative error properties; see Galbraith and Roberts, 2012). In other words, the dose estimator of the CAM generally does not converge towards the arithmetic mean and thus may lead to age underestimates, except in the presumably exceptional cases of symmetrical dose rates distributions (as may be the case in homogeneous environments from a radioactivity spatial distribution perspective). This demonstration is supported by empirical evidence from a set of well-behaved sediment samples (both in terms of luminescence characteristics and depositional history – i.e. unaffected by either post-depositional mixing or incomplete bleaching) for which independent chronological information is available (Guérin et al., 2015a; Thomsen et al., 2016; see also the discussion in Guérin et al., 2015b).

4.2. The Average Dose Model

The CAM appears to suffer from two main weaknesses: (i) all the overdispersion is treated as a measurement uncertainty, whereas we argue that only the intrinsic overdispersion should be so considered (see section 2 above); and (ii) the CAM dose estimator does not converge to the average dose absorbed by the grains.

[Insert Fig. 5 here]

With these considerations in mind, we propose an Average Dose Model for the estimation of the mean dose absorbed by an assembly of quartz grains subject to variable natural dose rates, so as to verify Eq. (5). Fig. 5 shows a hierarchical representation of the Average Dose Model. First, we write the relationship between the dose absorbed by grain \( j \) and its equivalent dose as

\[
D_{e,j} = \Delta_j e^{\varepsilon_j + \sigma_m \eta_j}
\]  

or, in logarithmic space,

\[
d_j = \delta_j + \varepsilon_j + \sigma_m \eta_j
\]

where \( \delta_j = \log (\Delta_j) \), \( d_j = \log (D_{e,j}) \), \( \sigma_m \) is the intrinsic overdispersion (e.g. determined by applying the CAM to a dose recovery experiment), \( \varepsilon_j \) is the analytical uncertainty as defined in Eq. (7) and \( \eta_j \) is a centred reduced Gaussian variable as in Eq. (8). Ideally, the \( \sigma_m \) parameter should be defined by a gamma dose recovery experiment, i.e. with grains all having received the same dose before loading.
into the luminescence reader; this would presumably reproduce as closely as possible a homogeneous irradiation in nature. However, easy access to gamma irradiations is not common in the dating community and a beta dose recovery measurement may be the only practical alternative; in the latter case it must be borne in mind that Thomsen et al. (2005) showed that such an experiment may lead to an underestimation of $\sigma_m$.

It should be noted here that in Eq. (17), it is assumed that $\sigma_m$ is common to all grains, in particular (i) irrespective of their luminescence sensitivity and (ii), since Thomsen et al. (2012) showed an OD dependency on dose, independently of how close to saturation the natural luminescence signal lies. The idea behind the latter assumption is that the intrinsic OD might increase as the natural luminescence signal gets closer to the OSL saturation level. We tested this assumption for two samples, the results can be found in Supplementary Material; in summary, our tests showed that the intrinsic OD neither significantly depends on luminescence sensitivity nor on how close to saturation the natural luminescence signal lies.

It is expected that dose rates to single grains most commonly (in fact, whenever radioactive hotspots, such as K-feldspar grains or heavy minerals like zircons are present in the sediment) follow lognormal distributions (Nathan et al., 2003; Mayya et al., 2006; Guérin et al., 2015b). As a result, we model the $\Delta_j$ as

$$\Delta_j \sim \log\mathcal{N}(\mu, \sigma_d).$$

(18)

where $\sigma_d$ is unknown and characterises the dispersion in single grain dose rates. Based on Eq. (6), we then have (see, e.g., Johnson et al., 1994):

$$\Delta = E(\Delta_j) = e^{\mu + \frac{\sigma_d^2}{2}}$$

(19)

which is equivalent to

$$\mu = \log(\Delta) - \frac{\sigma_d^2}{2}.$$  

(20)

Thus, the $\Delta_j$ are linked with $\Delta$ through

$$\Delta_j = \Delta e^{-\frac{\sigma_d^2}{2} e^{\sigma_d v_j}}$$

(21)

with $v_j \sim \mathcal{N}(0,1)$. To summarise,

$$D_{e,j} \sim \log\mathcal{N}(\log\Delta - \frac{\sigma_d^2}{2}, \sigma_m^2 + \sigma^2)$$

(22)

which can be rewritten in the form

$$D_{e,j} = \Delta e^{-\frac{\sigma_d^2}{2} e^{\sigma_m v_j} + \sigma_d v_j} \Leftrightarrow d_j = \log\Delta - \frac{\sigma_d^2}{2} + \varepsilon_j + \sigma_m v_j + \sigma_d v_j$$

(23)
The log-likelihood function is then given by

\[
\log L(d_1, \ldots, d_n, \sigma_{1, \ldots, \sigma_n}, \Delta, \sigma_d) = \sum_j -\log \left( \frac{2\pi}{\sigma_j^2 + \sigma_m^2 + \sigma_d^2} \right) - \frac{(d_j - \log \Delta + \frac{\sigma_d^2}{2})^2}{2(\sigma_j^2 + \sigma_m^2 + \sigma_d^2)}.
\]  

(24)

As a result, the maximum likelihood estimators of \( \Delta \) and \( \sigma_d \) verify the two following equations:

\[
\log \Delta = \frac{\sum_j d_j + \frac{\sigma_d^2}{2}}{\sum_j \frac{1}{\sigma_j^2 + \sigma_m^2 + \sigma_d^2}}
\]

(25)

and

\[
\sum_j \frac{1 + d_j - \log \Delta + \frac{\sigma_d^2}{2}}{\sigma_j^2 + \sigma_m^2 + \sigma_d^2} = \sum_j \frac{(d_j - \log \Delta + \frac{\sigma_d^2}{2})^2}{(\sigma_j^2 + \sigma_m^2 + \sigma_d^2)^2}
\]

(26)

It should be noted here that there is no explicit formula to calculate the maximum likelihood of \( \sigma_d \). Furthermore, there is a solution to Eq. (26) only if the left-hand side term is greater than or equal to zero. In that case, it is possible to compute the maximum likelihood estimators of \( \Delta \) and \( \sigma_d \). Otherwise, the log-likelihood function is decreasing and its maximum is reached when \( \sigma_d = 0 \).

An R (R development Core Team, 2016) script implementing the Average Dose Model is provided as supplement and available in the R ‘Luminescence’ package (Kreutzer et al., 2012). The standard errors on \( \Delta \) and \( \sigma_d \) are calculated by bootstrapping (Efron and Tibshirani, 1986).

5. Results and discussion

Guérin et al. (2015a) used a set of independently known age samples to test the accuracy of the CAM and of a central dose Bayesian model proposed by Combès et al. (2015). Here, we re-use the same samples, with the same data analysis in terms of grain selection and curve fitting of the dose response curves. The only difference concerns samples Bdx 16045 to 16049 (n=5) for which (i) multi-grain OSL has been measured and (ii) more grains are available for the single grain analysis (with only marginal changes in CAM results compared to those published by Guérin et al., 2015a). The comparison between the CAM and the ADM ages is straightforward, since both models require the same measurement data, i.e. lists of \( D_{e,j} \) and associated \( \sigma_j \) values. The only additional parameter needed to run the ADM is the additional measurement error \( \sigma_m \), determined as the intrinsic overdispersion – which is characteristic of both the analysed sample and the central dose to be determined (cf. Thomsen et al., 2012).

[Insert Table 3 here]

[Insert Fig. 6 here]
For each sample, we re-calculated the ages using the ADM. Fig. 6 shows ADM and CAM-based ages as a function of independent age (see also Table 3, which includes multi-grain OSL ages). Linear regression of the two data sets indicates that the ADM-based ages are systematically closer to the references than the standard CAM-based ages. Furthermore, we can compare the OSL to reference age ratios obtained with multi-grain, CAM-based single grain and ADM- based single-grain datasets: this ratio is 0.994 ± 0.024 for multi-grain ages (n=18), 0.925 ± 0.021 for CAM-based single grain ages and 0.987 ± 0.021 for ADM-based ages (n=19 in both latter cases). The first conclusion that can be drawn from these averages and standard errors is that CAM-based single grain ages are, on average, not consistent with independent age (see also Thomsen et al., 2016); the CAM appears to lead to age underestimations by on average 8 ± 2 %, even although such a systematic underestimation could not be predicted based on the average dose recovery ratio (Table 1). This result confirms the prediction of Ineq. (15) above. Secondly, a paired t-test on the OSL to reference age ratios shows that the CAM and ADM give statistically different results (p<0.001). Conversely, multi-grain and single-grain ADM-based ages are statistically consistent with each other (paired t-test, p=0.48) and both sets of ages are, on average, consistent (at two standard errors) with independent age control. It should be emphasised here that the agreement between multi-grain ages and single grain, ADM-based ages is very encouraging, given the overall reliability of multi-grain OSL ages as shown by several reviews in the literature (e.g. Murray and Olley, 2002; Rittenour, 2008).

In Fig. 7, the relative differences between the OSL ages calculated with the CAM dose and the reference ages is plotted as a function of the same quantity when the OSL age calculation is performed with the ADM dose. Points lying above the 1:1 line (13 out of 19) indicate that ADM ages are more accurate than CAM ages. Thus, both Figs. 6 and 7 confirm that ADM-based ages are more accurate than CAM-based ages.

In contrast, such complex $D_e$ probability distributions are taken into account in the Bayesian model of Combès et al. (2015), which leads us to hypothesise that the BaSAR model handles larger doses better compared to the CAM and ADM. Thus, it seems that while in general quartz OSL age underestimation is a widely acknowledged concern (e.g. Buylaert et al., 2007), at least part of it can be
attributed to inadequate data analysis (see also the discussion in Guérin et al., 2015b). However, further testing of this hypothesis is beyond the scope of the present study.

6. Implications for other age models and multi-grain OSL estimates

6.1. Consequences for the use of the MAM and FMM

A corollary of this discussion and results concerns other age models besides the CAM, since the dose estimation parameters of the Minimum Age Model (MAM; Galbraith and Laslett, 1993) and the Finite Mixture Model (FMM; Galbraith and Green, 1990) are the same as those of the CAM. Since it appears that the CAM must generally lead to age underestimates (Ineq.15), it follows that the MAM and FMM must also lead to age underestimates.

6.2. A further note on the use of the FMM

We now turn to the implications of this discussion on the use of the FMM. The FMM is mainly used to separate two – or more – discrete dose components in a sediment sample. These components are generally presumed to have resulted from the mixing of grains from two different layers of different ages (say, layers 1 and 2). While this assumption may be reasonable (although difficult to prove) when a stratigraphic record has been affected by post-depositional processes, the effect of such mixing has consequences on dose rates that have rarely been discussed. Deeben et al. (2013) noted that the dose rate experienced by the grains before mixing is unknown and may be different from today’s measured dose rate (i.e. after mixing); as a result, the authors advocated caution in the interpretation of FMM ages. To formalise the problem, if one assumes that the mixing of layers 1 and 2 occurred a time \( t_m \) ago, then we can write that:

\[
\Delta_1 = \overline{D}_1 (t_1 - t_m) + (f_1 \overline{D}_1 + f_2 \overline{D}_2) t_m
\]  

(27)

where \( t_1 \) is the age of sediment deposition of layer 1, and \( f_1 \) and \( f_2 \) represent the proportions of layer 1 and 2 in the mixing of these two layers (a similar equation can also be written for layer 2). This equation quite simply states that any mixing of sufficient magnitude to be reflected in the dose distribution must, in general, also have an impact on radioelement concentrations. We can rewrite Eq. (27) as:

\[
t_1 = t_m + \frac{\Delta_1 - (f_1 \overline{D}_1 + f_2 \overline{D}_2) t_m}{\overline{D}_1}
\]

(28)

since the aim of using the FMM is to determine \( t_1 \) (or \( t_2 \)). In this equation, the FMM may provide a (biased, i.e. underestimated) value of \( \Delta_1 \) and estimates of \( f_1 \) and \( f_2 \) (even although these estimates are likely to be significantly in error; cf. Roberts et al., 2000; Guérin et al., 2013). However, both \( t_m, \overline{D}_1 \) and \( \overline{D}_2 \) are unknown, (obviously one would have taken a sample from layer 1, and another from layer 2, had this been possible in the field). In other words, either the ingredients necessary for the age calculation are absent, or modelling could be avoided. As a consequence, we must regard published FMM ages as of doubtful value, except possibly in cases where dose rates do not vary significantly through the section containing the layers of interest.
The only exceptions to this rule are (i) if dose rates from the original Layers 1 and 2 were identical – but this assumption cannot be tested; or (ii) in the particular case where the originally upper layer was simultaneously deposited and mixed with the older layer. In such a case, the lowest dose component identified by the FMM could, in principle, be used in conjunction with the measured, mixed dose rate to calculate the age of the grains in the originally upper layer. It should be emphasised here that it is the dose value which identifies the grains that can be accurately dated by applying the FMM, rather than the proportion of grains in the various components. It is usual in the literature to select the component in which the majority of grains are found, but there is no a priori reason to expect that the measured dose rate to this component has applied throughout the burial period.

These issues may reflect the poor terminology used in our field: so-called age models (CAM, MAM, FMM) are, in fact, dose models; they do not consider dose rates. We agree with the suggestion (e.g., Bailiff et al., 2013) that one should refer to the Central Dose Model (CDM) instead of CAM (and, similarly, Minimum Dose Model instead of Minimum Age Model) since these models are generally applied to equivalent doses (as it was already noted by Galbraith and Roberts, 2012).

6.3. Multi-grain OSL age estimates

The importance of correct statistical analysis usually becomes greater as the dispersion in OSL data increases; for a given sample, this usually follows decreasing aliquot size (see, e.g., Cunningham et al., 2011). Nevertheless, the discussion above has relevance to central dose estimation in the case of multi-grain OSL dating. If the aliquot size becomes sufficiently large that (i) there is a negligibly small variation in (relative) analytical uncertainties determined for individual $D_e$ values, (ii) these uncertainties are negligible compared to the scatter in $D_e$ values, and (iii) the intrinsic overdispersion parameter $\sigma_{in}$ also is negligible compared to $\sigma_{di}$ it can be easily shown from Eq. (22) that the unweighted arithmetic mean (empirical average) of individual dose estimates will tend to $\Delta$. Thus, in such conditions, for the sake of simplicity we advocate the use of the empirical average of $D_e$ values when calculating the best estimate of the age of a sample from multi-grain aliquots $D_e$ measurements (it should be noted that in the present study, some multi-grain ages were calculated using the CDM – see Notes of Table 3 for details).

8. Conclusion

The definition of equivalent dose implies that if any overdispersion is observed in OSL dose recovery distributions, it can only result from unidentified measurement errors. Conversely, extrinsic sources of dispersion in natural $D_e$ distributions such as those arising from dose rate heterogeneities are of a very different nature. Here we have, for the first time in the statistical analysis of single grain OSL data, discussed the effect of experimentally-measured dose rate parameters on ages. We point out that the Central Age Model (as well as the MAM and FMM) estimates the median (or geometric mean) of a lognormal distribution, whereas parameters determined for the distribution of corresponding dose rates are averages (arithmetic means). As a result, CAM-based single grain OSL doses will inevitably underestimate the burial doses and thus result in age underestimation. This model prediction has been validated by a series of experiments on samples for which independent chronological information is available. The amount of age underestimation is, for our sample set, $8 \pm 2\%$. A new model, the Average Dose Model (ADM; code available in the R ‘Luminescence’ package) has been introduced to address the identified weaknesses of the CAM. The ADM can be applied to well-bleached samples and leads to more accurate ages, which in our study are, on average, in agreement with independent ages. However, for both models an increasing age underestimation is observed with increasing age; we attribute this, at least in part, to improper analytical treatment of the effect of saturation of the quartz OSL signal with dose.
Finally, we argue that more appropriate acronyms should be used in the literature, such as CDM (for Central Dose Model) instead of CAM (and similarly, MDM instead of MAM). Ages calculated using the FMM should only be used with caution, and preferably avoided altogether, because the average dose rates absorbed by the grains of the different components cannot be known, except in very specific cases. For multi-grain aliquots, dividing a simple unweighted arithmetic mean of individual $D_e$ values by the average dose rate is, in general, most likely to give accurate ages.

Acknowledgements.

The authors are grateful for the financial supports of Région Aquitaine (in particular through the CHRO QUI programme) and of the French National Research Agency via the LaScArBx Labex (Project number ANR-10-LABX-52, in particular through the MAM and COVADIM projects). The authors also would like to thank Jakob Wallinga and one anonymous reviewer for constructive and helpful comments on an earlier version of this article.

References


Figure captions.

Figure 1. Intrinsic OD (OD\textsubscript{int}) as a function of absorbed dose, for all samples studied here. Despite scatter in the data, gamma dosed samples (red circles) tend to exhibit greater OD\textsubscript{int} values than beta dosed samples (black squares). In addition, OD\textsubscript{int} seems to increase with dose (as already noted by Thomsen et al., 2012).

Figure 2. Radial plots showing the natural D\textsubscript{e} distribution for sample BR-2011-8. In Fig. 2a, only the analytical errors are included; however, since these uncertainties are, by definition, underestimates of the real uncertainties, we advocate the display of uncertainties including the intrinsic overdispersion (equal to 13% in this case; Fig. 2b) as determined in a dose recovery experiment. The latter radial plot more faithfully represents our knowledge of D\textsubscript{e} measurement uncertainties.

Figure 3. Hierarchical representation of the CAM. See section 4.1 for an explanation of the different variables.

Figure 4. Radial plot showing the natural D\textsubscript{e} distribution for sample BR-2011-8. The CAM OD has been added in quadrature to analytical uncertainties to provide the uncertainty associated with each D\textsubscript{e} estimate that is taken into account in the central dose value by the CAM (Eq. 13). It appears that all the more precise D\textsubscript{e} estimates have essentially the same uncertainty, dominated by the natural overdispersion, and are thus given the same weight in the weighted geometric mean calculation of the CAM.

Figure 5. Hierarchical representation of the Average Dose Model. See the text (section 4.2) for the explanation of the different variables.

Figure 6. Single grain OSL ages, calculated either with the CAM (blue triangles and corresponding linear fit: blue dashed line) or the ADM (red circles and corresponding linear fit: red dashed line) as a function of independent chronological information. ADM ages seem to be more accurate than CAM ages.

Figure 7. Relative difference between single-grain OSL CAM-based ages and reference ages, as a function of the same quantity when OSL ages are estimated with the ADM. The black line is the 1:1 line: points lying above the line indicate a better performance of the ADM (n=14); points below the line indicate a better performance of the CAM (n=5).

Figure 8. CAM- and ADM-based single-grain OSL to reference age ratio, as a function of reference age. In both cases, the slope of the fitted line is statistically different from zero and thus seems to indicate a loss of accuracy of OSL ages with increasing age.
BR-2011-8

Analytical errors only

Standardised estimate vs. Equivalent dose [Gy]

Relative error [%]

Precision
BR-2011-8

Errors including intrinsic OD

Equivalent dose [Gy]

Sandared estimate

Relative error [%]

100  50  33.3  25  20  16.7  14.3

0   1   2   3   4   5   6   7

Precision
\[ j \in \{1, \ldots, n\} \]
BR-2011-8

Errors including CAM OD

Standardised estimate

Equivalent dose [Gy]

Relative error [%]

Precision

200 100 66.7

0.0 0.5 1.0 1.5
\( j \in \{1, \ldots, n\} \)

\[ \sigma_d \quad \Delta \quad \sigma_m \quad \Delta_j \quad \sigma_j \quad D_{e,j} \]

\[ j \in \{1, \ldots, n\} \]
The graph shows a scatter plot comparing OSL ages (ka) against age control (ka). The data points are color-coded and labeled as follows:

- Blue triangles: SG OSL, CAM
- Red circles: SG OSL, ADM

The solid black line represents the 1:1 relationship (y = x), while the dashed lines show the linear regression fits:

- Dotted blue line: y = 0.963 x
- Dotted red line: y = 0.921 x

The graph indicates a strong correlation between OSL ages and age control, with the majority of data points clustering close to the 1:1 line.
The graph shows the relationship between OSL/Reference age ratio and age in ka (kinematic ages). The data points are distinguished by color: red for ADM and blue for CAM.

The equations for the linear fits are:

- For ADM:
  \[ y = 1.01 (\pm 0.03) - 0.0022 (\pm 0.0011) x \]

- For CAM:
  \[ y = 0.97 (\pm 0.03) - 0.0025 (\pm 0.0011) x \]
<table>
<thead>
<tr>
<th>Sample</th>
<th>Given dose (Gy)</th>
<th>Radiation type</th>
<th>Meas. to given dose</th>
<th>σ</th>
<th>OD&lt;sub&gt;int&lt;/sub&gt; (%)</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>092201</td>
<td>107</td>
<td>gamma</td>
<td>0.88</td>
<td>0.03</td>
<td>29</td>
<td>3</td>
</tr>
<tr>
<td>092202</td>
<td>107</td>
<td>gamma</td>
<td>1.04</td>
<td>0.02</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>092203</td>
<td>74.8</td>
<td>gamma</td>
<td>0.89</td>
<td>0.02</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>092204</td>
<td>74.8</td>
<td>gamma</td>
<td>0.91</td>
<td>0.03</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>FER 3</td>
<td>56.8</td>
<td>gamma</td>
<td>1.07</td>
<td>0.03</td>
<td>17</td>
<td>3</td>
</tr>
<tr>
<td>InterComp</td>
<td>4.81</td>
<td>gamma</td>
<td>1.00</td>
<td>0.03</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>GDB 5</td>
<td>18.4</td>
<td>beta</td>
<td>0.99</td>
<td>0.02</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>GDB 3</td>
<td>73.4</td>
<td>beta</td>
<td>0.95</td>
<td>0.04</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>EBC 1</td>
<td>32.1</td>
<td>beta</td>
<td>1.02</td>
<td>0.02</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>EBC 2</td>
<td>41.3</td>
<td>beta</td>
<td>1.04</td>
<td>0.02</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Bdx 16045</td>
<td>9.60</td>
<td>beta</td>
<td>1.08</td>
<td>0.03</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>Bdx 16046</td>
<td>9.60</td>
<td>beta</td>
<td>1.02</td>
<td>0.02</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Bdx 16047</td>
<td>9.60</td>
<td>beta</td>
<td>1.03</td>
<td>0.02</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Bdx 16048</td>
<td>9.60</td>
<td>beta</td>
<td>1.05</td>
<td>0.02</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Bdx 16049</td>
<td>9.60</td>
<td>beta</td>
<td>1.05</td>
<td>0.01</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>BR2011-32</td>
<td>22.5</td>
<td>beta</td>
<td>0.98</td>
<td>0.02</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>BR-2011-8</td>
<td>11.0</td>
<td>beta</td>
<td>0.96</td>
<td>0.02</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>BR2011-15</td>
<td>22.5</td>
<td>beta</td>
<td>0.90</td>
<td>0.03</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>BR-2011-11</td>
<td>22.0</td>
<td>beta</td>
<td>0.98</td>
<td>0.03</td>
<td>15</td>
<td>3</td>
</tr>
</tbody>
</table>

Mean: 0.99 0.01

**Table 1.** Dose recovery data. List of samples and associated given doses (for details on the samples, see Guérin et al., 2015b). Beta doses were delivered with a $^{90}$Sr/$^{90}$Y source in the luminescence reader. Gamma doses were delivered with a reference $^{137}$Cs source in Risø DTU. ‘σ’ denotes the standard error on the preceding column. The CAM was used to calculate measured to given dose ratios and associated uncertainties, as well the intrinsic OD (OD<sub>int</sub>) values. We regard this overdispersion as characterizing un-recognized measurement errors. These values are used as input for the Mean Dose Model, in which they are called $σ_m$. 


<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t [ka] )</td>
<td>Sample age</td>
</tr>
<tr>
<td>( \Delta_j ) [Gy]</td>
<td>Dose absorbed by grain ( j )</td>
</tr>
<tr>
<td>( D_j ) [Gy.ka(^{-1})]</td>
<td>Dose rate to which grain ( j ) was exposed</td>
</tr>
<tr>
<td>( \bar{D} ) [Gy.ka(^{-1})]</td>
<td>Average dose rate, which is also the experimentally derived dose rate</td>
</tr>
<tr>
<td>( \Delta ) [Gy]</td>
<td>Sample palaeodose</td>
</tr>
<tr>
<td>( D_{e,j} ) [Gy]</td>
<td>Equivalent dose for grain ( j )</td>
</tr>
<tr>
<td>( \delta_j ) [Gy]</td>
<td>Relative analytical uncertainty on ( D_{e,j} )</td>
</tr>
<tr>
<td>( \gamma_j ) [Gy]</td>
<td>“true” logged equivalent dose for grain ( j )</td>
</tr>
<tr>
<td>( D_{e,CAM} ) [Gy]</td>
<td>CAM dose</td>
</tr>
<tr>
<td>( d )</td>
<td>Log ((D_{e,CAM}))</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>CAM overdispersion</td>
</tr>
<tr>
<td>( \sigma_m )</td>
<td>Additional measurement uncertainty (in practice, intrinsic overdispersion)</td>
</tr>
<tr>
<td>( \sigma_d )</td>
<td>Relative standard deviation in single grain dose rates</td>
</tr>
</tbody>
</table>

Table 2. List of main variables.
Table 3. Multi grain and Single grain OSL ages in comparison with independent age information (for more details see Guérin et al., 2015b). 'Ref. age' and 'pseudo-sigma' stand for the age obtained independently from OSL measurements. 'o' denotes the standard error on the preceding column. 'n' is the number of grains analysed for each sample. 'Age ratio' corresponds to the ratio of OSL to reference ages. 'OD (%)' corresponds to the overdispersion, estimated with the CAM, of the natural D<sub>e</sub> distribution. 'ADM' corresponds to ages calculated using the Average Dose Model.
Notes: for some samples, due to simultaneous publications and differences in data treatment (e.g., dose response curve fitting) the values quoted here may slightly differ from those mentioned in publications dedicated to the studied sites. For sample 092204, the age reference was calculated assuming that one of the $^{14}$C samples from Thomsen et al. (2016) is an outlier (Beta-234193 was removed from the analysis). Multi-grain ages were calculated using an unweighted average of D$_e$ values for samples 092201-04, FER3 and Intercomp; using the CDM for samples GDB 53, EBC 1 and 2, Bdx 16045-49 and BR-2011-8, -11, -15 and -32.
Is $\sigma_m$ is common to all grains?

Thomsen et al. (2012) showed, for both a heated and a bleached samples given a 250 mGy dose, that there is a dependency of OD on brightness – the running mean of OD decreasing with increasing brightness. Furthermore, they also showed that the (relative) OD increases with increasing dose, which might be related to curve fitting or more precisely that the OD increases when the natural signal approach the OSL saturation level. This might mean that our assumption that the intrinsic overdispersion is the same for all grains is not valid – thus, we tested the assumption that OD neither depends on brightness nor on where the natural signal lies on the dose response curve.

To check these assumptions, we used two datasets: first, calibration quartz allowed us to have a large dataset with highly variable sensitivities. We sorted the grains by increasing response to the first test dose (as a proxy for sensitivity) and separated them in deciles. In Fig. 1, the OD is plotted as a function of the decile number, called brightness index; it appears that, except for the first decile, i.e. for the 10% dimmest grains where the OD is significantly larger than for the whole population (n=452 grains), there is no relationship between OD and brightness. This result is consistent with the decrease of the OD, observed by Thomsen et al. (2012: Fig. 5b) when the number of grains included in the calculation, after sorting them by increasing brightness, is increased.

![Fig. S1. Intrinsic OD (%) as a function of brightness for calibration quartz.](image)

To test the dependency of OD on where the natural signal lies on the dose response curve (i.e. how close to the saturation level the natural signal is), we used a sample (TA2255) from the same site as samples 092201-04, for which a large number of grains (n=907) were subject to a high dose (180 Gy) beta dose recovery test. The dose response curves were fitted using saturating exponential functions so that we would better see effects linked with saturation, and we sorted the grains by ascending $D_e/D_o$. Here again, we then separated the grains in deciles and Fig. 2 shows the OD variations as a function of $D_e/D_o$. It appears that, despite rather important fluctuations, there is no systematic relationship between OD and $D_e/D_o$. As a result, we conclude that our assumption that the same intrinsic OD can be added to all grains is valid.

---

For such tests to be performed, we need large datasets, in particular larger than the samples measured in this study; as a result, we used samples already measured independently of the present study. We assume that the fundamental properties investigated here apply to all samples.
**Fig. S2.** Intrinsic OD (%) as a function of $D_e/D_0$ (the higher this ration, the closer the natural signal lies to the saturation level of the dose response curve).