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

Guillaume Guérin, Georges Rossi, A. Philippe, S. Murray, K.J. Thomsen, Chantal Tribolo, Petra Urbanová, M. Jain, Pierre Guibert, Norbert Mercier, et al.

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

G. Guérin[1,*], C. Christophe[1], A. Philippe[2], A.S. Murray[3], K.J. Thomsen[4], C. Tribolo[1], P. Urbanova[1], M. Jain[4], P. Guibert[1], N. Mercier[1], S. Kreutzer[1], C. Lahaye[1]

[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:

$$t = \frac{\Delta}{\dot{D}}, \quad (1)$$

where t is the age (in ka), Δ is the absorbed dose (in Gy) and $\dot{D}$ 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_{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_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_{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_e distribution, Fig. 2a is clearly misleading.

{Insert Fig. 2 here}

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_{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_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_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_e distributions in order to identify appropriate statistical analysis. We regard the intrinsic overdispersion parameter, OD_{int}, as describing unrecognised measurement errors in D_e; this definition now allows us to move on to discuss the relationship between the D_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_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* ($\forall j$)

$$\forall j,\ t = \frac{\Delta_j}{\dot{D}_j}, \tag{2}$$

where Δ_j is the dose absorbed by grain *j* and $\dot{D}_j$ is the dose rate to which this grain has been exposed. For clarity, Table 2 lists all variables used in the following equations.

{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.*, Al_2O_3: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}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}U, ^{235}U, and ^{232}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}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:

$$\overline{\dot{D}} = \frac{\sum_j \dot{D}_j}{n}. \quad (3)$$

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

$$\forall j,\ \Delta_j = t\dot{D}_j \Leftrightarrow \frac{\sum_j \Delta_j}{n} = t\frac{\sum_j \dot{D}_j}{n}, \quad (4)$$

Then,

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

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

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

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 (Δ) 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 ($D_{e,j}$, σ_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 σ 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 $d_j = log(D_{e,j})$, δ_j is the logged "true" equivalent dose, ε_j is an error term such that $\mathbb{E}(\varepsilon_j) = 0$ and $var(\varepsilon_j) = \sigma_j^2$, and

$$\delta_j = d + \sigma\eta_j \tag{8}$$

where $d = log(D_{e,CAM})$ 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 $D_{e,CAM}$ and σ. Eqs. (7) and (8) can be compacted as:

$$d_j = d + \varepsilon_j + \sigma\eta_j \tag{9}$$

or

$$D_{e,j} = D_{e,CAM} e^{\varepsilon_j + \sigma\eta_j}. \tag{10}$$

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

$$D_{e,j} \sim log\mathcal{N}(d, \sigma^2 + \sigma_j^2). \tag{11}$$

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

$$\log L(D_{e,1},\ldots,D_{e,n}, \sigma_1,\ldots,\sigma_n, d, \sigma) = \sum_j -\log\left(\sqrt{2\pi}\sqrt{\sigma^2 + \sigma_j^2}\right) - \frac{(d_j - d)^2}{2\left(\sigma_j^2 + \sigma^2\right)}. \tag{12}$$

Thus, the maximum likelihood estimators $\hat{d}$ and $\hat{\sigma}$ of *d* and σ verify the following equations in *d* and σ:

$$d = \frac{\sum_j \frac{d_j}{\sigma^2 + \sigma_j^2}}{\sum_j \frac{1}{\sigma^2 + \sigma_j^2}} \tag{13}$$

and

$$\sum_j \frac{1}{\sigma^2 + \sigma_j^2} = \sum_j \frac{(d_j - d)^2}{(\sigma^2 + \sigma_j^2)^2}\ . \tag{14}$$

Since the exponential function is regular, $e^{\hat{d}}$ is the maximum likelihood estimator of $e^d = D_{e,CAM}$, 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 (σ_j) 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

$$t \geq \frac{D_{e,CAM}}{\bar{D}} \ . \tag{15}$$

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 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

$$D_{e,j} = \Delta_j e^{\varepsilon_j + \sigma_m \eta_j} \tag{16}$$

or, in logarithmic space,

$$d_j = \delta_j + \varepsilon_j + \sigma_m \eta_j \tag{17}$$

where $\delta_j = \log(\Delta_j)$, $d_j = \log(D_{e,j})$, σ_m is the intrinsic overdispersion (*e.g.* determined by applying the CAM to a dose recovery experiment), ε_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 σ_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 σ_m.

It should be noted here that in Eq. (17), it is assumed that σ_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 Δ_j as

$$\Delta_j \sim log\mathcal{N}(\mu,\sigma_d). \tag{18}$$

where σ_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):

$$\Delta = \mathbb{E}(\Delta_j) = e^{\mu + \frac{\sigma_d^2}{2}} \tag{19}$$

which is equivalent to

$$\mu = \log(\Delta) - \frac{\sigma_d^2}{2}. \tag{20}$$

Thus, the Δ_j are linked with Δ through

$$\Delta_j = \Delta e^{-\frac{\sigma_d^2}{2}} e^{\sigma_d v_j} \tag{21}$$

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

$$D_{e,j} \sim log\mathcal{N}(log\Delta - \frac{\sigma_d^2}{2}, \sigma_j^2 + \sigma_m^2 + \sigma_d^2) \tag{22}$$

which can be rewritten in the form

$$D_{e,j} = \Delta e^{-\frac{\sigma_d^2}{2}} e^{\varepsilon_j + \sigma_m \eta_j + \sigma_d v_j} \Leftrightarrow d_j = log\Delta - \frac{\sigma_d^2}{2} + \varepsilon_j + \sigma_m \eta_j + \sigma_d v_j. \tag{23}$$

335 The log-likelihood function is then given by

336 $$\log L(d_1,\ldots,d_n, \sigma_1,\ldots,\sigma_n,\Delta,\sigma_d) = \sum_j -\log\left(\sqrt{2\pi(\sigma_j^2+\sigma_m^2+\sigma_d^2)}\right) - \frac{\left(d_j - log\Delta + \frac{\sigma_d^2}{2}\right)^2}{2\left(\sigma_j^2+\sigma_m^2+\sigma_d^2\right)} \quad (24)$$

337 As a result, the maximum likelihood estimators of Δ and σ_d verify the two following equations:

338 $$\log\Delta = \frac{\sum_j \frac{d_j + \frac{\sigma_d^2}{2}}{\sigma_j^2+\sigma_m^2+\sigma_d^2}}{\sum_j \frac{1}{\sigma_j^2+\sigma_m^2+\sigma_d^2}} \quad (25)$$

339 and

340 $$\sum_j \frac{1 + d_j - log\Delta + \frac{\sigma_d^2}{2}}{\sigma_j^2+\sigma_m^2+\sigma_d^2} = \sum_j \frac{(d_j - log\Delta + \frac{\sigma_d^2}{2})^2}{(\sigma_j^2+\sigma_m^2+\sigma_d^2)^2} \quad (26)$$

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

345 An R (R development Core Team, 2016) script implementing the Average Dose Model is
346 provided as supplement and available in the R 'Luminescence' package (Kreutzer *et al.*, 2012). The
347 standard errors on Δ and σ_d are calculated by bootstrapping (Efron and Tibshirani, 1986).

348 **5. Results and discussion**

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

360 ***{Insert Table 3 here}***

361 ***{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.

{Insert Fig. 7 here}

In Fig. 8, the ratio of single grain OSL to reference age is plotted as a function of reference age. It appears that the accuracy of the OSL ages decreases with age (the slope of the fitted line is -2.2 ± 1.1 10^{-3} ka^{-1} for ADM-based ages and -2.5 ± 1.0 10^{-3} ka^{-1} for CAM-based ages and thus significantly differs, in the latter case, from zero). This trend was already observed for CAM-based ages by Guérin *et al.* (2015a; their Fig. 5b). So, while the accuracy of ADM-based ages is generally better than that of CAM-based ages, a loss of accuracy seems to be associated with increasing age. Conversely, the Bayesian model BaSAR of Combès *et al.* (2015) did not show such a trend (the slope of the line obtained by Guérin *et al.*, 2015a, is 0.2 ± 1.1 10^{-3} ka^{-1} – cf. their discussion and Fig. 5a). This might be explained by the fact that the CAM and ADM require lists of $D_{e,j}$ and σ_j values, *i.e.* simple parameterisations of individual equivalent data: the probability density of the D_e of each grain/aliquot is described by a lognormal distribution (Eqs. (9) and (16)). However, when the natural signal lies on a non-linear portion of the dose response curve (*e.g.*, close to saturation for grains having a near-zero linear component), the variance in the probability density of D_e values becomes increasingly large and the lognormal distribution may not satisfactorily describe this density. At present, we cannot think of a simple function that would describe, better than lognormal distributions, both aliquots in the linear range of the dose response curve and aliquots close to saturation. A way around this is provided, *e.g.* in the Analyst software, by the use of Monte Carlo simulations of both L_n/T_n and dose response curves (Berger, 2010; Duller, 2015). In such a case, the D_e probability density distributions are more complex; but these distributions cannot be fed into the ADM (nor into the CAM) as these are simple parametric statistical models.

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.

{Insert Fig. 8 here}

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:

$$\Delta_1 = \overline{\dot{D}}_1(t_1 - t_m) + (f_1\overline{\dot{D}}_1 + f_2\overline{\dot{D}}_2)t_m \quad (27)$$

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:

$$t_1 = t_m + \frac{\Delta_1 - (f_1\overline{\dot{D}}_1 + f_2\overline{\dot{D}}_2)t_m}{\overline{\dot{D}}_1} \quad (28)$$

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 Δ_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, $\overline{\dot{D}}_1$ and $\overline{\dot{D}}_2$ 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 $\sigma_m$also is negligible compared to σ_d, it can be easily shown from Eq. (22) that the unweighted arithmetic mean (empirical average) of individual dose estimates will tend to Δ. 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 ± 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 CHROQUI 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

Aitken, M. J., 1985. Thermoluminescence dating. Academic Press, London.

Arnold, L.J., Roberts, R.G., Galbraith, R.F., DeLong, S.B., 2009. A revised burial dose estimation procedure for optical dating of young and modern-age sediments. Quaternary Geochronology, 4, 306-325.

Bailiff, I.K., Lewis, S.G., Drinkall, H.C., White, M.J., 2013. Luminescence dating of sediments from a Palaeolithic site associated with a solution feature on the North Downs of Kent, UK. Quaternary Geochronology 18, 135-148.

Banerjee, D., 2001. Supralinearity and sensitivity changes in optically stimulated luminescence of annealed quartz. Radiation Measurements 33, 47-57.

Berger, G.W., 2010. Estimating the error in equivalent dose values obtained from SAR. Ancient TL 28, 55–66.

Combès, B., Lanos, P., Philippe, A., Mercier, N., Tribolo, C., Guérin, G., Guibert, P., Lahaye, C., 2015. A Bayesian central equivalent dose model for optically stimulated luminescence dating. Quaternary Geochronology 28, 62-70.

Cunningham, A.C., Wallinga, J. and Minderhoud, P.S.J. 2011. Expectations of scatter in equivalent-dose distributions when using multi-grain aliquots for OSL dating. Geochronometria 38, 424-431.
Cunningham, A.C., Wallinga, J. and Minderhoud, P.S.J. 2011. Expectations of scatter in equivalent-dose distributions when using multi-grain aliquots for OSL dating. Geochronometria 38, 424-431.

Deeben, J., Hiddink, H., Huisman, D.J., Müller, Schokker, J., Wallinga, J. 2010. Middle Palaeolithic artefact migration due to periglacial processes; a geological investigation into near-surface occurrence of Palaeolithic artefacts (Limburg – Eastern Brabant coversand region, the Netherlands). Netherlands Journal of Geosciences – Geologie en Mijnbouw 89, 35-50.

Dietze, M., Kreutzer, S., Burow, C., Fuchs, M. C., Fischer, M., Schmidt, C., 2016. The abanico plot: visualising chronometric data with individual standard errors. Quaternary Geochronology 31, 12-18.

Duller, G.A.T., 2015. The Analyst software package for luminescence data: overview and recent improvements. Ancient TL 33, 35-42.

Efron, B., Tibshirani, R., 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Statistical science, 54-75.

Galbraith, R.F., 1988. Graphical Display of Estimates Having Differing Standard Errors. Technometrics 30, 271–281.

Galbraith, R., 2015. On the mis-use of mathematics: A comment on “How confident are we about the chronology of the transition between Howieson's Poort and Still Bay?” by Guérin *et al.* (2013). Journal of Human Evolution, 80, 184-186.

Galbraith, R., Green, P., 1990. Estimating the component ages in a finite mixture. International Journal of Radiation Applications and Instrumentation. Part D. Nuclear Tracks and Radiation Measurements 17, 197–206.

Galbraith, R.F., Laslett, G.M., 1993. Statistical models for mixed fission track ages. Nuclear Tracks And Radiation Measurements 21, 459–470.

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

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

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

Guérin, G., Mercier, N. and Adamiec, G, 2011. Dose-rate conversion factors: update. Ancient TL, 29 (1), 5-8.

Guérin, G., B. Combès, C. Tribolo, C. Lahaye, N. Mercier, P. Guibert, K. J. Thomsen, 2015a. Testing the accuracy of a single grain OSL Bayesian central dose model with known-age samples. Radiation Measurements, 81, 62-70.

Guérin, G., Jain M., Thomsen K. J., Murray A. S., Mercier, N., 2015b. Modelling dose rate to single grains of quartz in well-sorted sand samples: the dispersion arising from the presence of potassium feldspars and implications for single grain OSL dating. Quaternary Geochronology 27, 52-65.

Guérin, G., Frouin, M., Talamo, S., Aldeias, V., Bruxelles, L., Chiotti, L., Dibble, H. L., Goldberg, P., Hublin, J.-J., Jain, M., Lahaye, C., Madelaine, S., Maureille, B, McPherron, S. P., Mercier, N., Murray, A. S., Sandgathe, D., Steele, T. E., Thomsen, K. J., Turq, A., 2015c. A Multi-method Luminescence Dating of the Palaeolithic Sequence of La Ferrassie Based on New Excavations Adjacent to the La Ferrassie 1 and 2 Skeletons, au Journal of Archaeological Science 58, 147-166.

Guérin, G., Frouin, M., Tuquoi, J., Thomsen, K.J., Goldberg, P., Aldeias, V., Lahaye, C., Mercier, N., Guibert, P., Jain, M., Sandgathe, D., McPherron, S.J.P., Turq, A., Dibble, H.L., 2016. The complementarity of luminescence dating methods illustrated on the Mousterian sequence of the Roc de Marsal: A series of reindeer-dominated, Quina Mousterian layers dated to MIS 3. Quaternary International, in press.

Huntley, D.J., 2001. Note: some notes on language. Ancient TL 19, 27–28.

Huntley, D. J., Godfrey-Smith, D. I., Thewalt, M. L. W., 1985, Optical dating of sediments, Nature, 313, 105–7.

Hütt, G., Jaek, I., Tchonka, J., 1988. Optical dating: K-feldspars optical response stimulation spectra, Quaternary Science Reviews, 7, 381–5.

ICRP, 2007. The 2007 Recommendations of the International Commission on Radiological Protection. ICRP Publication 103. Ann. ICRP 37 (2-4).

Jain, M., Bøtter-Jensen, L., Thomsen, K.J., 2007. High local ionization density effects in x-ray excitations deduced from optical stimulation of trapped charge in Al2O3:C. Journal of Physics: Condensed Matter, 19, 116-201.

Johnson, N. L., Kotz, S., Balakrishnan, N., 1994. 14: Lognormal Distributions, Continuous univariate distributions. Vol. 1, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics (2nd ed.). New York: John Wiley & Sons.

Kreutzer, S., Schmidt, C., Fuchs, M.C., Dietze, M., Fischer, M., Fuchs, M., 2012. Introducing an R package for luminescence dating analysis. Ancient TL 30, 1–8.

Mayya, Y.S., Morthekai, P., Murari, M.K., Singhvi, A.K., 2006. Towards quantifying beta microdosimetric effects in single-grain quartz dose distribution. Radiation Measurements, 41, 1032-1039.

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

Murray, A.S., Olley, J.M., 2002. Precision and Accuracy in the Optically Stimulated Luminescence Dating of Sedimentary Quartz: A Status Review. Geochronometria 21, 1-16.

Murray, A.S., Buylaert, J.-P., Thiel, C., 2015. A luminescence dating intercomparison based on a Danish Beach-ridge sand. Radiation Measurements 81, 32-38.

Nathan, R.P., Thomas, P.J., Jain, M., Murray, A.S., Rhodes, E.J., 2003. Environmental dose rate heterogeneity of beta radiation and its implications for luminescence dating: Monte Carlo modelling and experimental validation. Radiation Measurements 37, 305-313.

Olley, J.M., Roberts, R.G., Murray, A.S., 1997. Disequilibria in the uranium decay series in sedimentary deposits at Allen's cave, nullarbor plain, Australia: implications for dose rate determinations. Radiation Measurements 27, 433–443.

Pawitan, Y., 2001. In all likelihood: statistical modelling and inference using likelihood. Oxford University Press.

R Development Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://cran-project.org

Reimann, T., Lindhorst, S., Thomsen, K.J., Murray, A.S., Hass, C.H., Frechen, M., 2012. OSL dating of mixed coastal sediments from Sylt (German Bight, North Sea). Quaternary Geochronology 11, 52-67.

Rittenour, T.M., 2008. Luminescence dating of fluvial deposits: applications to geomorphic, palaeoseismic and archaeological research. Boreas 37, 613-635.

Roberts, R.G., Galbraith, R.F., Yoshida, H., Laslett, G.M., Olley, J.M., 2000. Distinguishing dose populations in sediment mixtures: a test of single-grain optical dating procedures using mixtures of laboratory-dosed quartz. Radiation Measurements 32, 459-465.

612 Thomsen K. J., Murray A. S., Bøtter-Jensen L., 2005. Sources of variability in OSL dose measurements
613 using single grains of quartz. Radiation Measurements, 39, 47-61.

614 Thomsen K. J., Murray A. S., Jain M., 2012. The dose dependency of the over-dispersion of quartz OSL
615 single grain dose distributions. Radiation Measurements, 47, 732-739.

616 Thomsen, K.J., Murray, A.S., Buylaert, J.-P., Jain, M., Helt-Hansen, J., Aubry, T., 2016. Testing single-
617 grain quartz OSL methods using known age samples from the Bordes-Fitte rockshelter (Roches
618 d'Abilly site, Central France). Quaternary Geochronology 31, 77-96.

619 Vandenberghe, D., De Corte, F., Buylaert, J. P., & Kučera, J., 2008. On the internal radioactivity in
620 quartz. Radiation Measurements, 43, 771-775.

621

622

Figure captions.

Figure 1. Intrinsic OD (OD_{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_{int} values than beta dosed samples (black squares). In addition, OD_{int} seems to increase with dose (as already noted by Thomsen *et al.*, 2012).

Figure 2. Radial plots showing the natural D_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_e measurement uncertainties.

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

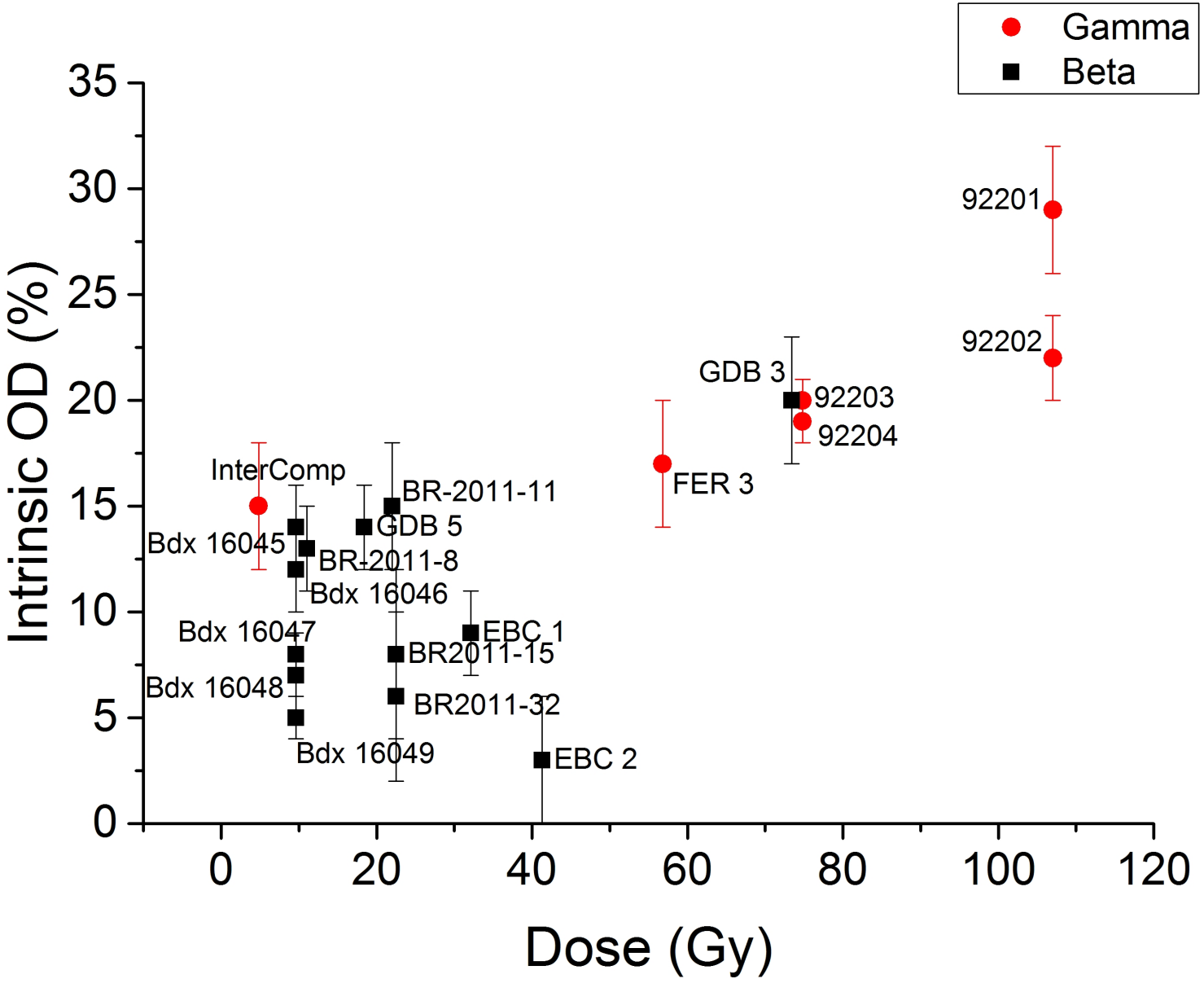
Figure 4. Radial plot showing the natural D_e distribution for sample BR-2011-8. The CAM OD has been added in quadrature to analytical uncertainties to provide the uncertainty associated with each D_e estimate that is taken into account in the central dose value by the CAM (Eq. 13). It appears that all the more precise D_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.

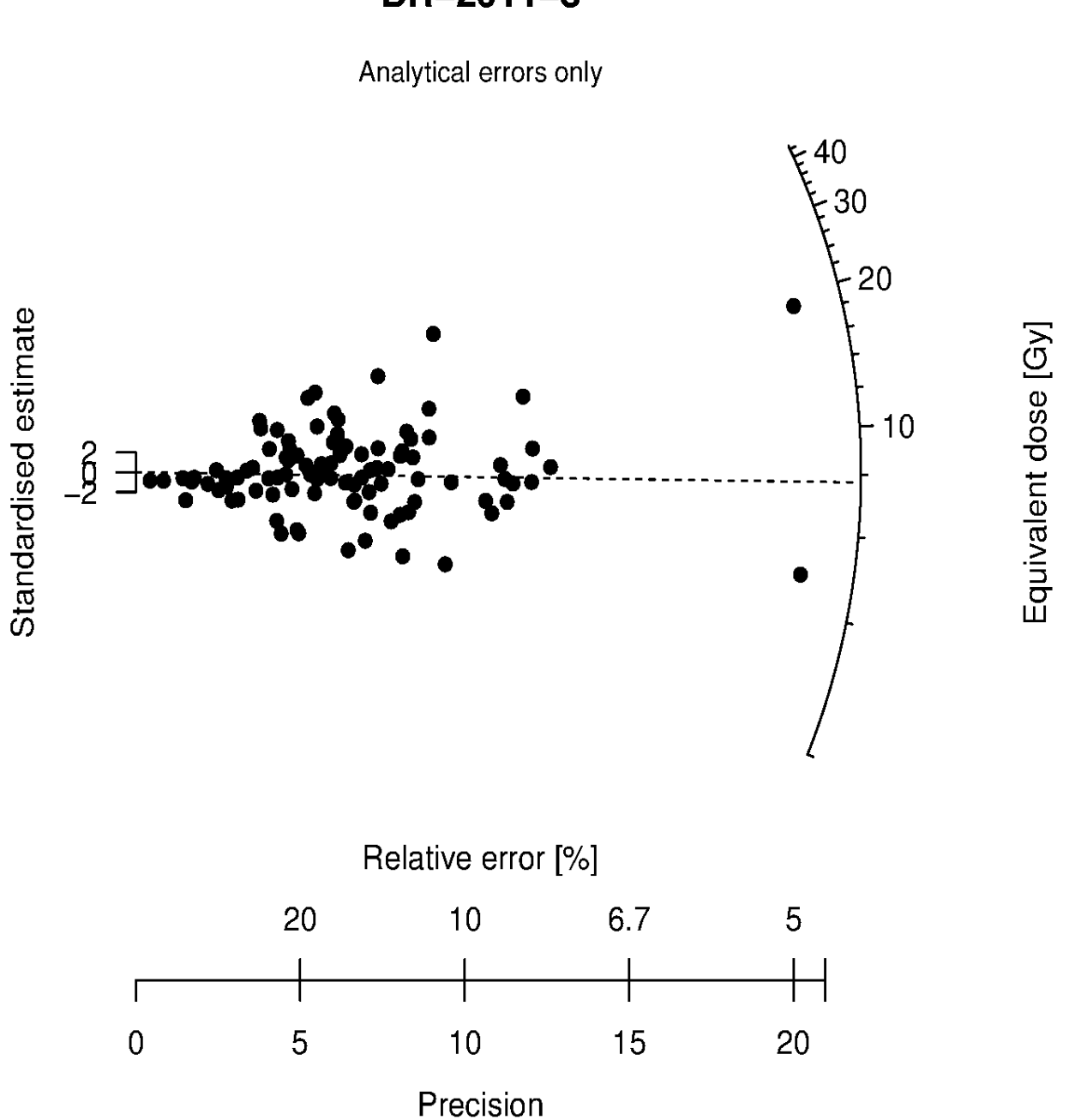
Figure 5. Hierarchical representation of the Average Dose Model. See the text (section 4.2) for the explanation of the different variables.

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.

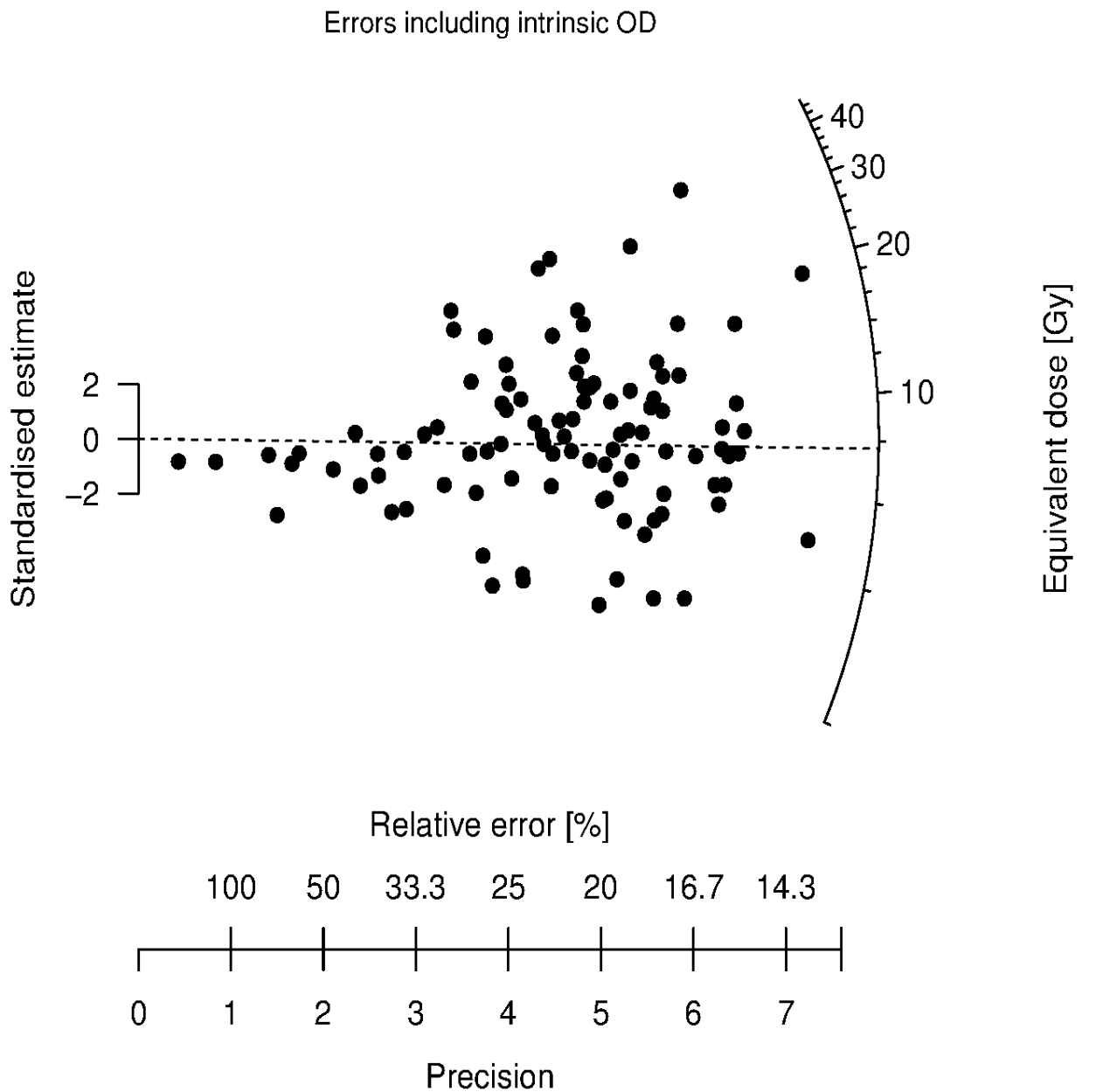
Figure 7. Relative difference between single-grain OSL CAM-based ages and reference ages, as a function of the same quantity when OSL ages are estimated with the ADM. The black line is the 1:1 line: points lying above the line indicate a better performance of the ADM (n=14); points below the line indicate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

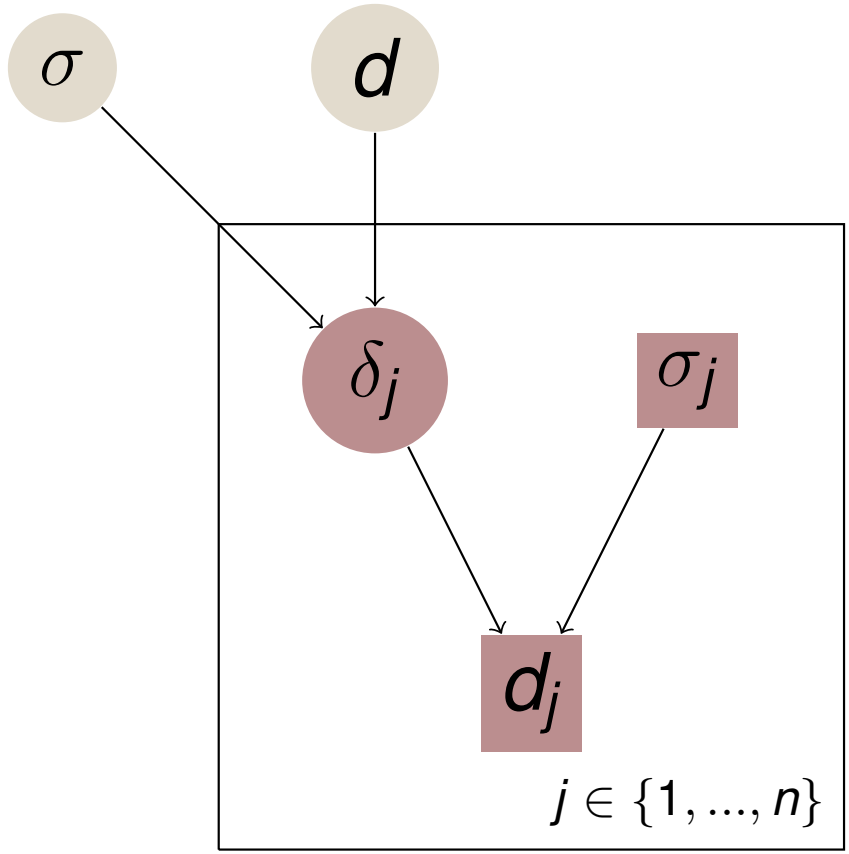
Figure 8. CAM- and ADM-based single-grain OSL to reference age ratio, as a function of reference age. In both cases, the slope of the fitted line is statistically different from zero and thus seems to indicate a loss of accuracy of OSL ages with increasing age.

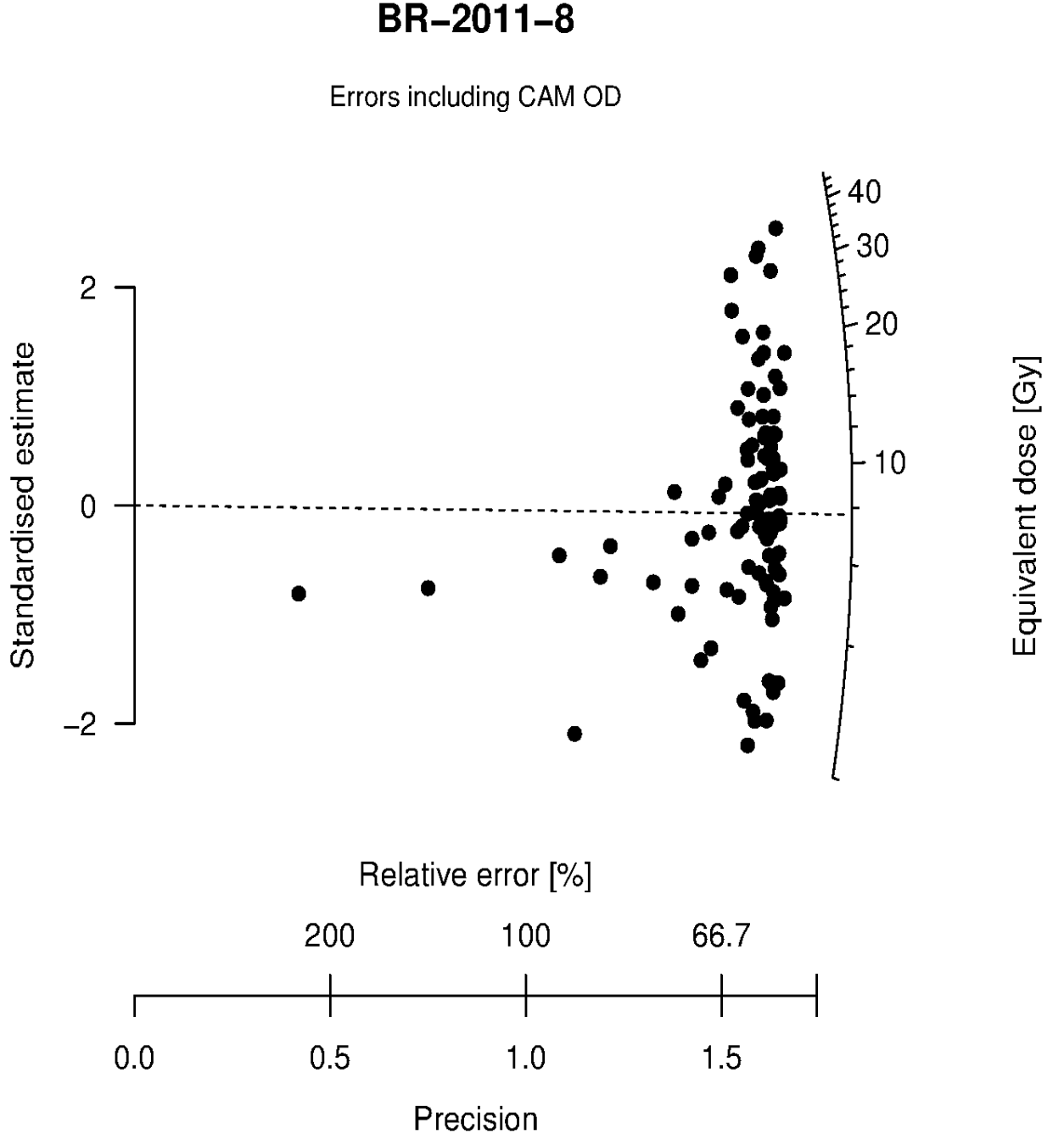

Gamma
Beta
Intrinsic OD (%)
Dose (Gy)
0
5
10
15
20
25
30
35
0
20
40
60
80
100
120
92201
92202
GDB 3
92203
92204
FER 3
InterComp
BR-2011-11
GDB 5
Bdx 16045
BR-2011-8
Bdx 16046
Bdx 16047
EBC 1
BR2011-15
Bdx 16048
BR2011-32
Bdx 16049
EBC 2

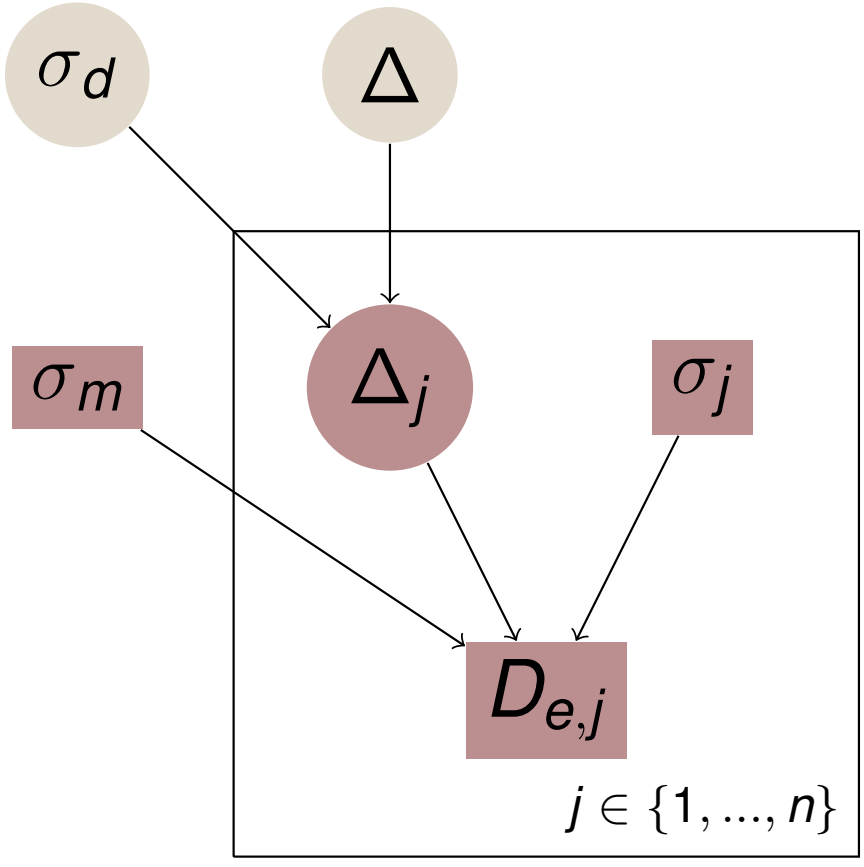

BR–2011–8
Analytical errors only
Standardised estimate
Equivalent dose [Gy]
40
30
20
10
Relative error [%]
20
10
6.7
5
0
5
10
15
20
Precision

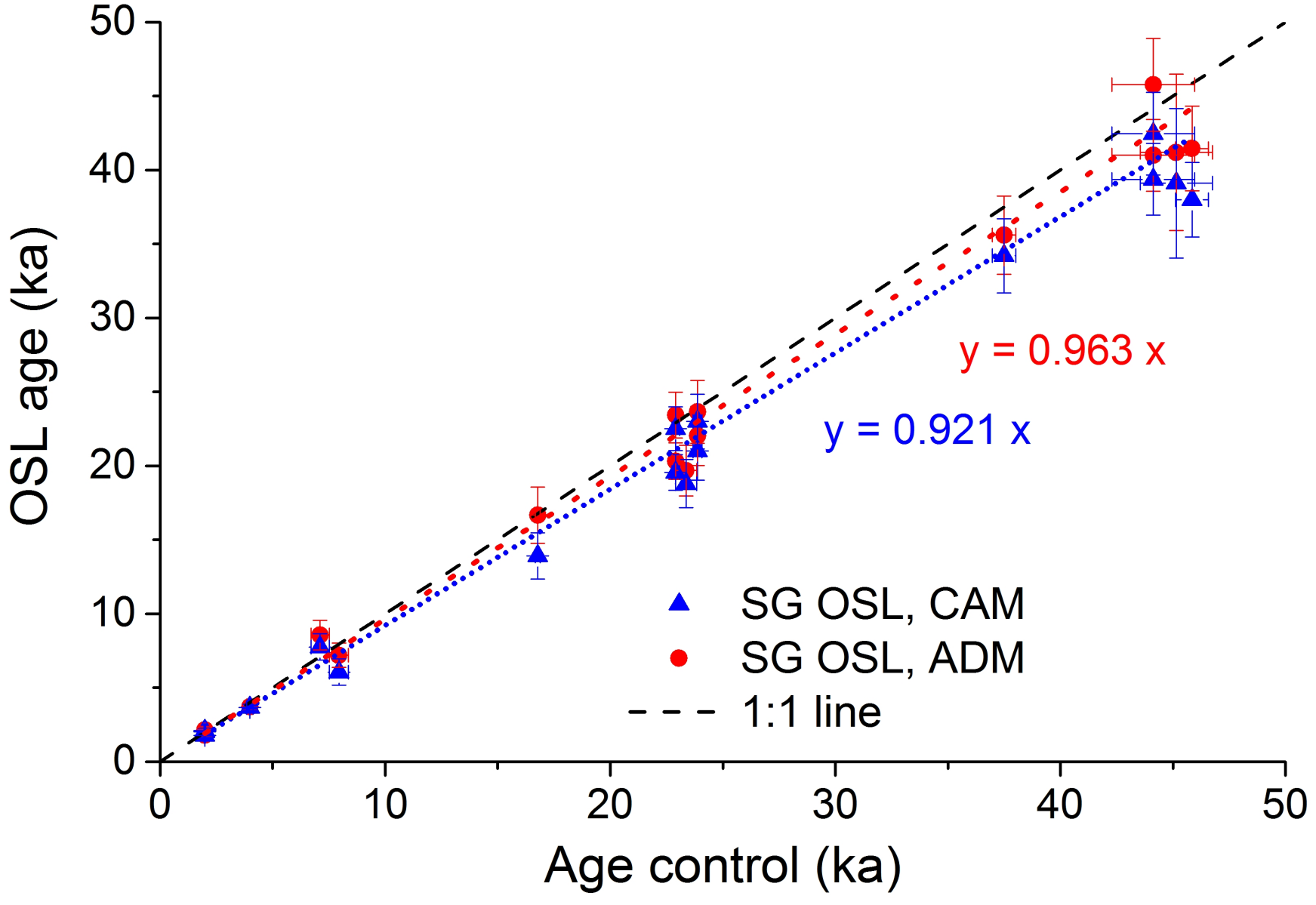
BR–2011–8

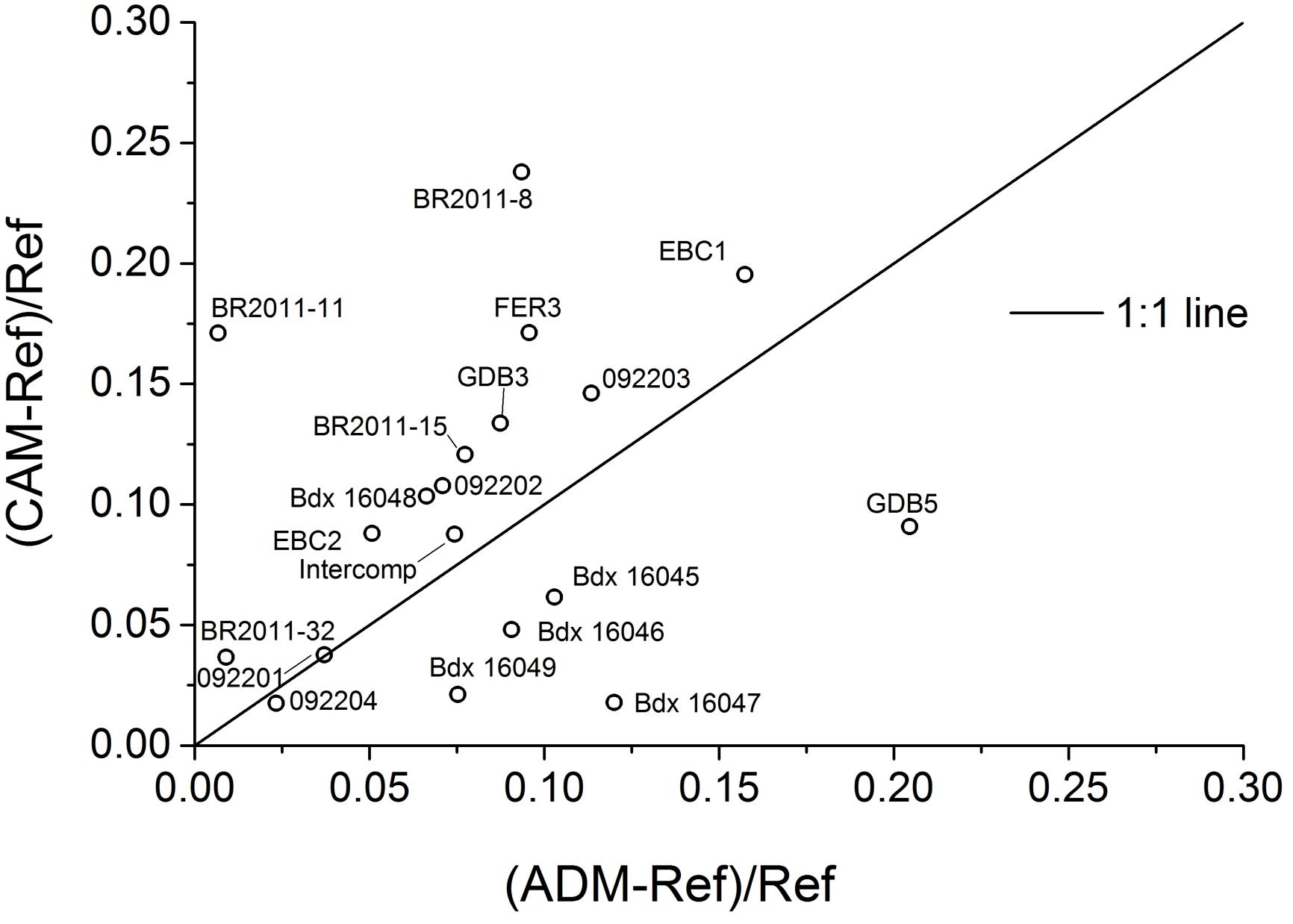

Errors including intrinsic OD
Standardised estimate
2
0
−2
40
30
20
10
Equivalent dose [Gy]
Relative error [%]
100
50
33.3
25
20
16.7
14.3
0
1
2
3
4
5
6
7
Precision

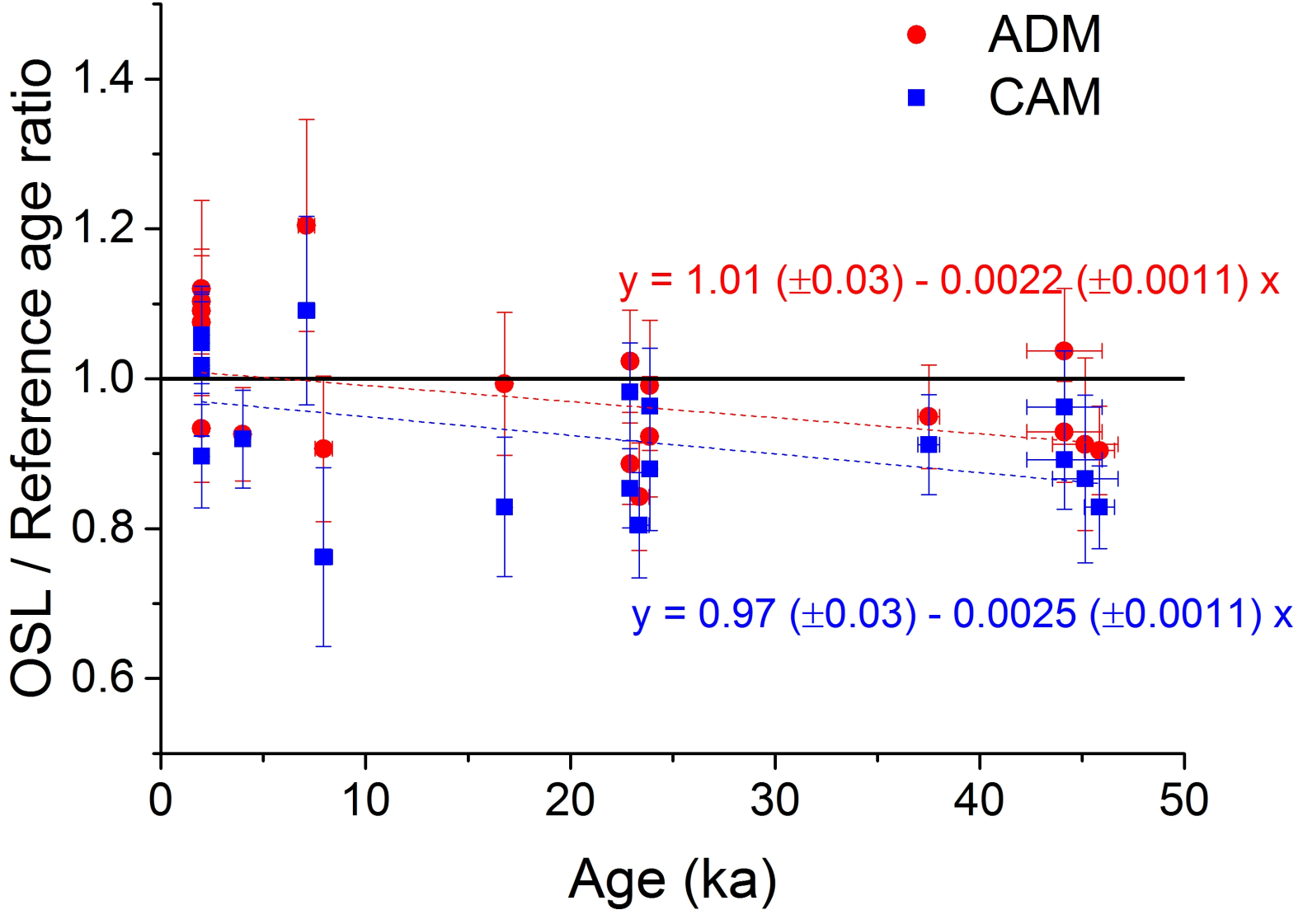

σ
d
δ_j
σ_j
d_j
$j \in \{1, ..., n\}$


BR–2011–8
Errors including CAM OD
Standardised estimate
2
0
−2
40
30
20
10
Equivalent dose [Gy]
Relative error [%]
200
100
66.7
0.0
0.5
1.0
1.5
Precision


σ_d
Δ
Δ_j
σ_m
σ_j
$D_{e,j}$
$j \in \{1, ..., n\}$


OSL age (ka)
Age control (ka)
0
10
20
30
40
50
y = 0.963 x
y = 0.921 x
SG OSL, CAM
SG OSL, ADM
1:1 line


0.30
0.25
0.20
0.15
0.10
0.05
0.00
(CAM-Ref)/Ref
0.00
0.05
0.10
0.15
0.20
0.25
0.30
(ADM-Ref)/Ref
BR2011-8
EBC1
BR2011-11
FER3
1:1 line
GDB3
092203
BR2011-15
092202
Bdx 16048
GDB5
EBC2
Intercomp
Bdx 16045
BR2011-32
Bdx 16046
092201
Bdx 16049
092204
Bdx 16047


ADM
CAM
y = 1.01 (±0.03) - 0.0022 (±0.0011) x
y = 0.97 (±0.03) - 0.0025 (±0.0011) x
OSL / Reference age ratio
1.4
1.2
1.0
0.8
0.6
0
10
20
30
40
50
Age (ka)

Sample	Given dose (Gy)	Radiation type	Meas. to given dose	σ	OD_{int} (%)	σ
092201	107	gamma	0.88	0.03	29	3
092202	107	gamma	1.04	0.02	22	2
092203	74.8	gamma	0.89	0.02	20	1
092204	74.8	gamma	0.91	0.03	19	1
FER 3	56.8	gamma	1.07	0.03	17	3
InterComp	4.81	gamma	1.00	0.03	15	3
GDB 5	18.4	beta	0.99	0.02	14	2
GDB 3	73.4	beta	0.95	0.04	20	3
EBC 1	32.1	beta	1.02	0.02	9	2
EBC 2	41.3	beta	1.04	0.02	3	3
Bdx 16045	9.60	beta	1.08	0.03	14	2
Bdx 16046	9.60	beta	1.02	0.02	12	2
Bdx 16047	9.60	beta	1.03	0.02	8	1
Bdx 16048	9.60	beta	1.05	0.02	7	1
Bdx 16049	9.60	beta	1.05	0.01	5	1
BR2011-32	22.5	beta	0.98	0.02	6	4
BR-2011-8	11.0	beta	0.96	0.02	13	2
BR2011-15	22.5	beta	0.90	0.03	8	4
BR-2011-11	22.0	beta	0.98	0.03	15	3
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_{int}) 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.

t *[ka]*	Sample age
Δ_j *[Gy]*	Dose absorbed by grain *j*
$\dot{D}_j$ *[Gy.ka^{-1}]*	Dose rate to which grain *j* was exposed
$\overline{\dot{D}}$ *[Gy.ka^{-1}]*	Average dose rate, which is also the experimentally derived dose rate
Δ *[Gy]*	Sample palaeodose
$D_{e,j}$ *[Gy]*	Equivalent dose for grain *j*
d_j	Log ($D_{e,j}$)
σ_j	Relative analytical uncertainty on $D_{e,j}$
δ_j	"true" logged equivalent dose for grain *j*
$D_{e,CAM}$ *[Gy]*	CAM dose
d	Log ($D_{e,CAM}$)
σ	CAM overdispersion
σ_m	Additional measurement uncertainty (in practice, intrinsic overdispersion)
σ_d	Relative standard deviation in single grain dose rates

Table 2. List of main variables.

					Multi grain				Single grain										
									CAM							ADM			
Sample	Ref. age (ka)	Pseudo-sigma	Dose rate	σ	Age (ka)	σ	Age ratio	σ	n	Age (ka)	σ	OD (%)	σ	Age ratio	σ	Age (ka)	σ	Age ratio	σ
092201	44.1	1.8	2.36	0.12	46.1	2.2	1.04	0.05	273	42.5	2.8	48	3	0.96	0.07	45.8	3.1	1.04	0.08
092202	44.1	1.8	2.51	0.12	42.4	2.2	0.96	0.05	218	39.4	2.4	35	3	0.89	0.06	41.0	2.4	0.93	0.07
092203	22.9	0.2	3.23	0.16	22.4	1.2	0.98	0.05	218	19.6	1.2	36	2	0.85	0.05	20.3	1.2	0.89	0.05
092204	22.9	0.2	2.47	0.13	28.0	1.4	1.22	0.06	146	22.5	1.5	34	3	0.98	0.06	23.4	1.5	1.02	0.07
FER 3	45.8	0.7	1.58	0.08	42.2	2.8	0.92	0.06	190	38.0	2.5	45	3	0.83	0.06	41.5	2.9	0.90	0.07
InterComp	3.99	0.14	1.24	0.06	3.99	0.14	1.00	0.04	123	3.67	0.23	30	3	0.91	0.07	3.69	0.22	0.95	0.06
GDB 5	7.11	0.41	2.51	0.24	-	-	-	-	189	7.75	0.89	49	3	1.09	0.14	8.56	1.00	1.20	0.15
GDB 3	45.1	1.6	2.26	0.25	37.4	5.3	0.83	0.12	101	39.1	5.1	39	4	0.87	0.12	41.2	5.3	0.92	0.12
EBC 1	23.4	0.5	1.62	0.11	19.1	1.7	0.82	0.07	129	18.8	1.6	24	2	0.80	0.07	19.7	1.7	0.82	0.07
EBC 2	37.5	0.5	1.42	0.07	36.6	2.8	0.98	0.08	198	34.2	2.5	21	1	0.91	0.07	35.6	2.6	0.93	0.07
Bdx 16045	1.99	0.04	1.83	0.07	2.07	0.14	1.04	0.07	196	2.10	0.12	31	2	1.06	0.07	2.19	0.13	1.10	0.06
Bdx 16046	1.99	0.04	1.72	0.08	2.14	0.15	1.08	0.08	141	2.08	0.13	30	2	1.05	0.07	2.16	0.14	1.09	0.07
Bdx 16047	1.99	0.04	2.29	0.17	1.99	0.16	1.00	0.08	139	2.02	0.18	44	3	1.02	0.09	2.22	0.23	1.12	0.09
Bdx 16048	1.99	0.04	2.07	0.13	1.94	0.14	0.97	0.07	165	1.78	0.13	29	2	0.90	0.07	1.85	0.14	0.93	0.07
Bdx 16049	1.99	0.04	1.92	0.14	1.72	0.13	0.87	0.07	119	2.01	0.17	36	2	1.02	0.09	2.13	0.19	1.08	0.09
BR-2011-32	23.9	0.2	1.08	0.08	24.8	2.7	1.04	0.11	125	23.0	1.9	30	3	0.96	0.08	23.7	2.1	1.01	0.09
BR-2011-8	7.94	0.43	1.34	0.12	7.91	0.80	1.00	0.10	99	6.05	0.89	60	5	0.76	0.12	7.20	0.83	0.91	0.15
BR-2011-15	23.9	0.2	1.04	0.08	23.8	2.62	1.00	0.11	68	21.0	1.95	33	4	0.88	0.08	22.0	2.0	0.92	0.10
BR-2011-11	16.8	0.1	1.02	0.07	19.3	2.0	1.15	0.12	62	13.9	1.6	62	6	0.83	0.09	16.7	1.9	1.00	0.15
Mean							0.976	0.021						0.925	0.021			0.987	0.022

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. 'σ' 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_e 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 σ_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[1]: 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.

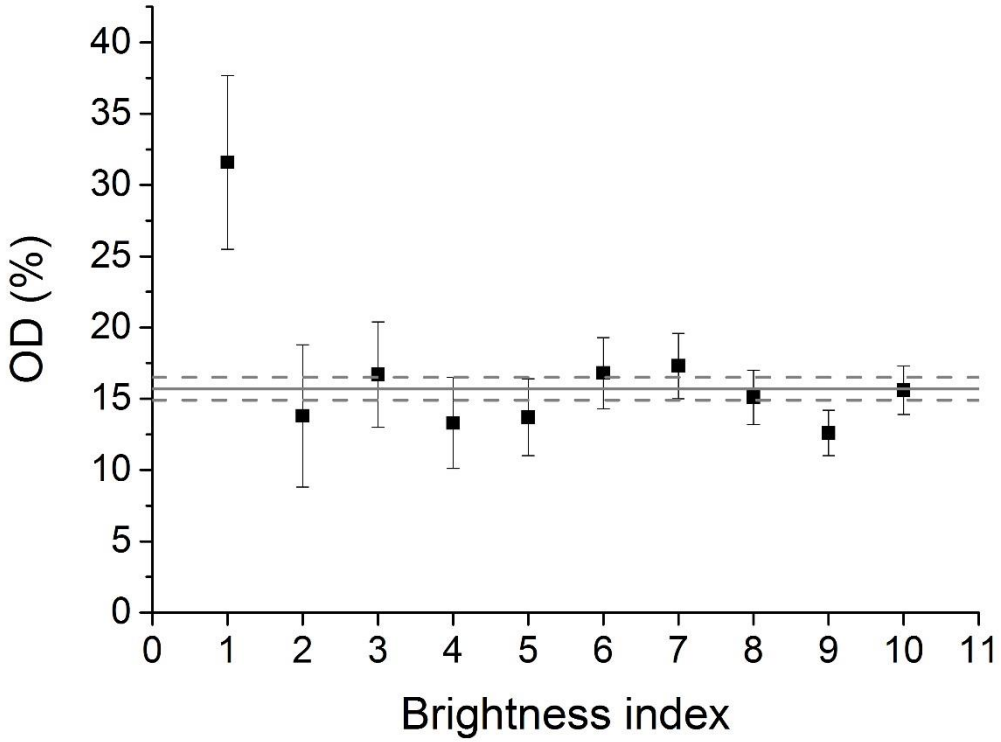


Fig. S1. Intrinsic OD (%) as a function of brightness for calibration quartz.

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_0. Here again, we then separated the grains in deciles and Fig. 2 shows the OD variations as a function of D_e/D_0. It appears that, despite rather important fluctuations, there is no systematic relationship between OD and D_e/D_0. As a result, we conclude that our assumption that the same intrinsic OD can be added to all grains is valid.

[1] 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.

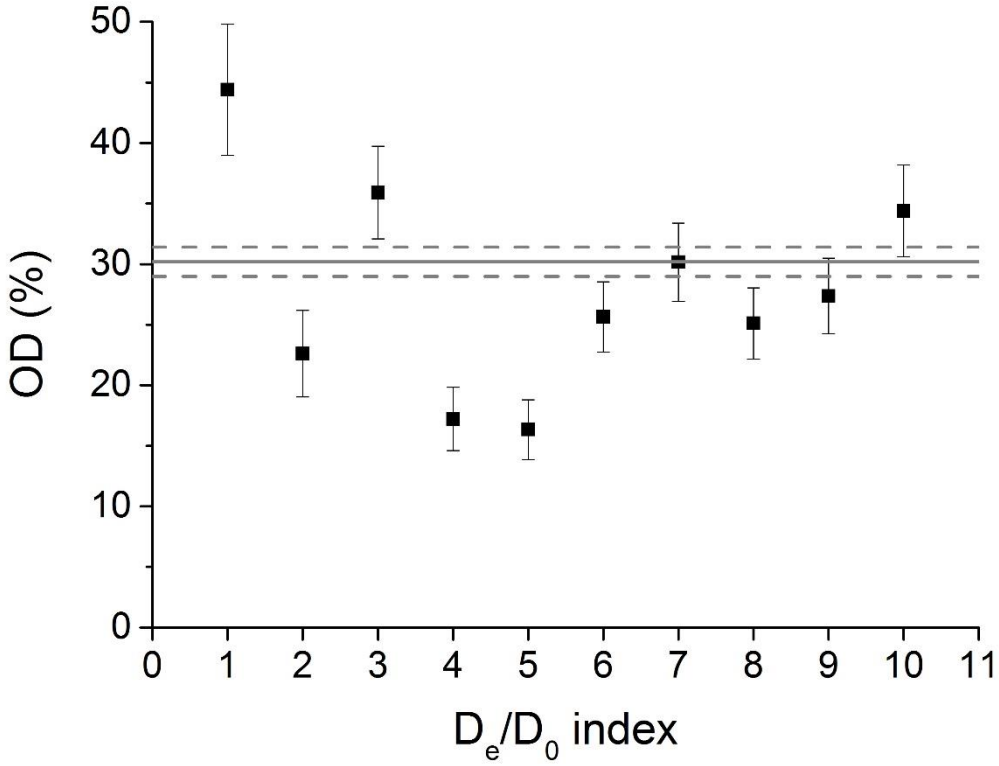


Fig. S2. Intrinsic OD (%) as a function of D_e/D_0 (the higher this ration, the closer the natural signal lies to the saturation level of the dose response curve).