



**HAL**  
open science

## 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.

### ► To cite this version:

Guillaume Guérin, Georges Rossi, A. Philippe, S. Murray, K.J. Thomsen, et al.. Absorbed dose, equivalent dose, measured dose rates, and implications for OSL age estimates: Introducing the Average Dose Model. *Quaternary Geochronology*, 2017, 41, pp.163 - 173. 10.1016/j.quageo.2017.04.002 . hal-01841513

**HAL Id: hal-01841513**

**<https://hal.science/hal-01841513v1>**

Submitted on 28 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

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

5 <sup>1</sup> UMR 5060 CNRS - Université Bordeaux Montaigne, IRAMAT-CRP2A, Maison de l'archéologie,  
6 Esplanade des Antilles, 33607 Pessac cedex, France

7 <sup>2</sup> Jean Leray Laboratory of Mathematics (LMJL), UMR6629 CNRS - Université de Nantes, France.

8 <sup>3</sup> Nordic Laboratory for Luminescence Dating, Department of Geoscience, Aarhus University, DTU Risø  
9 Campus, DK-4000 Roskilde, Denmark.

10 <sup>4</sup> Center for Nuclear Technologies, Technical University of Denmark, DTU Risø Campus, DK-4000  
11 Roskilde, Denmark.

12  
13  
14 **Abstract**

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

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

30 **Highlights:**

- 31 - Dose rate estimates correspond to arithmetic means  
32 - OSL age models should thus aim at arithmetic means of absorbed doses  
33 - We introduce the Average Dose Model (ADM)  
34 - ADM ages for known-age samples are more accurate than CAM ages  
35 - We argue against the use of the FMM  
36

37 **1. Introduction**

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

42 
$$t = \frac{\Delta}{\dot{D}}, \quad (1)$$

43 where  $t$  is the age (in ka),  $\Delta$  is the absorbed dose (in Gy) and  $\dot{D}$  is the dose rate (in Gy.ka<sup>-1</sup>).

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

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

62 **2. The concept of equivalent dose**

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

75 We note in passing that there is no alpha dose contribution to burial dose in multi- or single-  
76 grain dating when the grains have been etched to remove the outer alpha-irradiated layer. This is  
77 standard practice in all coarse-grain quartz dating; as a result, in the following we do not consider alpha

78 irradiations (except from a generally small internal dose rate contribution, following e.g.  
79 Vandenberghe *et al.*, 2008).

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

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

100 **{Insert Table 1 and Fig. 1 here}**

101 Based on this definition of equivalent dose (defined as the beta or gamma laboratory dose that  
102 results in the same signal intensity as the natural signal), we can now review the evidence concerning  
103 analytical uncertainties on luminescence-based  $D_e$  values published in the literature. These  
104 uncertainties, often estimated with the Analyst software (Duller, 2015), usually include contributions  
105 from counting statistics, curve fitting, and instrument reproducibility. The rate of absorption of dose  
106 by individual grains in nature is unknown; only a mass-averaged value is derived from dosimetry  
107 measurements (see section 3 below). As a result, at the single grain level one can only discuss an  
108 experimental comparison between equivalent dose and absorbed dose in the case of dose recovery  
109 experiments (Thomsen *et al.*, 2005). Table 1 lists the samples analysed (these samples were already  
110 analysed in Guérin *et al.*, 2015b; for details, the reader is referred to this publication) and the results  
111 of dose recovery experiments, including the intrinsic overdispersion ( $OD_{int}$ , following the terminology  
112 of Thomsen *et al.*, 2005). In such experiments, the latent OSL signals from samples are reset in some  
113 manner, usually by exposure to sufficient light, before a known dose is given under controlled  
114 laboratory conditions. Experimentally observed distributions of  $D_e$  are then compared to this known  
115 dose. These distributions usually display greater dispersion than that explained by the quantifiable  
116 analytical uncertainties, *i.e.* non-zero overdispersion (Table 1, Fig. 1; see also e.g. Galbraith *et al.*, 2005;  
117 Thomsen *et al.*, 2005, 2012). If the known dose was absorbed from gamma radiation, all grains can be  
118 experimentally arranged to have absorbed the same dose, and so any overdispersion must arise from  
119 the measurement process. In beta dose recovery experiments, where one is comparing beta  $D_e$  values  
120 with a given grain-specific beta dose (which does not need to be well known in absolute terms, *i.e.* in  
121 Gy), grains are exposed to the same electron spectra, at the same location on the disc, resting in the  
122 same position on the sample disc, etc., both for the given dose and for the regeneration doses; as a  
123 result many or most of the potential sources of scatter are avoided. Then any overdispersion in the  $D_e$

124 distribution is presumed to arise from unrecognised sources of measurement uncertainty. This  $OD_{int}$  is  
125 often non-trivial compared to quantifiable sources of uncertainty, and we must deduce that, in the  
126 literature, analytical uncertainties attributed to experimental  $D_e$  values are systematically  
127 underestimated, often by a large amount. This observation argues against the use of any graphical  
128 representation based on analytical uncertainties, whether radial plots (Galbraith, 1988) or abanico  
129 plots (Dietze *et al.*, 2015); since the true measurement uncertainties are unknown, such plots display  
130 false information and should be avoided. If they must be used, then presumably best estimates of  
131 uncertainty (*i.e.*, including  $OD_{int}$ ) should be used (as for example in Reimann *et al.*, 2012; Guérin *et al.*,  
132 2016). To illustrate this point, Fig. 2 shows two different radial plots for sample BR-2011-8 (Lahaye *et al.*,  
133 2015); in Fig. 2a only the analytical errors are taken into account, while in Fig. 2b the intrinsic  
134 overdispersion, which is by definition an additional measurement uncertainty, is added in quadrature  
135 to the individual analytical errors. If visual interpretation is to be used for the analysis of this single  
136 grain  $D_e$  distribution, Fig. 2a is clearly misleading.

137 **{Insert Fig. 2 here}**

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

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

### 161 **3. Measurements of dose rates**

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

167 With this in mind, let us consider a sediment sample including  $n$  identical quartz grains. In  
168 practice, Eq. (1) cannot directly be used for each of these grains because the dose rate to each quartz

169 grain is experimentally inaccessible. Therefore, assuming that all grains have the same age  $t$ , we can  
170 write based on Eq. (1) that for all  $j$  ( $\forall j$ )

$$171 \quad \forall j, t = \frac{\Delta_j}{\dot{D}_j}, \quad (2)$$

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

174 **{Insert Table 2 here}**

175  
176 Different methods are used for dose rate determination. Among the most frequently used are  
177 gamma spectrometry (both high resolution, laboratory-based, and *in situ* using portable scintillators),  
178 artificial dosimeters (e.g., Al<sub>2</sub>O<sub>3</sub>:C pellets), alpha and beta counting techniques, etc. (for a summary,  
179 see e.g. Aitken, 1985). In the more commonly used nuclide-specific techniques (e.g., neutron activation  
180 analysis, gamma spectrometry) the concentrations (or activities per unit mass) of various members of  
181 the U and Th decay chains, and of <sup>40</sup>K are measured directly. These concentrations are then used to  
182 calculate the infinite matrix dose rate using dose rate conversion factors. The latter provide the  
183 average energy emitted per disintegration, and are derived from tables of nuclear data (e.g., Guérin *et*  
184 *al.*, 2011). In the case of U- and Th-series, the average energy emitted per unit time and mass is  
185 summed over all daughters from <sup>238</sup>U, <sup>235</sup>U, and <sup>232</sup>Th, respectively. In other words, we estimate the  
186 total amount of energy emitted in the sample per unit time and unit mass. Using the infinite matrix  
187 assumption (Roesch and Attix, 1968; Aitken, 1985), this rate is equal to the rate of energy absorption  
188 per time and mass. For simplicity, let us assume that we have a sediment entirely made up of  $n$  identical  
189 grains of quartz, all having the same mass and age  $A$ , and that the radioactive sources have negligible  
190 mass. Since dose rate is the total amount of energy absorbed per unit time and mass in the sample,  
191 and since energy is a cumulative quantity, if one grain receives more than the average dose rate, say a  
192 fraction  $(1+x)$  of the average, then the remaining  $(n-1)$  grains receive, on average, a fraction  $(1-x)/(n-1)$   
193 of the average dose rate. If another grain receives a fraction  $(1-x)$  of the average dose rate, all  
194 remaining  $(n-2)$  grains must receive the average. This statement can be generalised: no matter the  
195 distribution of dose rates to individual grains, the invariant parameter is the amount of energy  
196 available for the grains, independent of how the radioactivity is distributed in the sample (see Guérin  
197 *et al.*, 2012b; Guérin *et al.*, 2015b).

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

209 Alpha, beta or gamma integral counting methods induce an additional systematic source of  
210 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:

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

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,

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

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

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

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 ( $\Delta$ ) 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

$$\delta_j = d + \sigma\eta_j \quad (8)$$

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

$$d_j = d + \varepsilon_j + \sigma\eta_j \quad (9)$$

or

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

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

$$D_{e,j} \sim \text{logN}(d, \sigma^2 + \sigma_j^2). \quad (11)$$

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:

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

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

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

and

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

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 ( $\sigma_j$ ) and the intrinsic overdispersion are negligible compared to the dispersion in dose rates to single grains, *i.e.* the

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

$$277 \quad t \geq \frac{D_{e,CAM}}{\bar{D}} . \quad (15)$$

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

#### 288 4.2. The Average Dose Model

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

293 **{Insert Fig. 5 here}**

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

$$298 \quad D_{e,j} = \Delta_j e^{\varepsilon_j + \sigma_m \eta_j} \quad (16)$$

299 or, in logarithmic space,

$$300 \quad d_j = \delta_j + \varepsilon_j + \sigma_m \eta_j \quad (17)$$

301 where  $\delta_j = \log(\Delta_j)$ ,  $d_j = \log(D_{e,j})$ ,  $\sigma_m$  is the intrinsic overdispersion (*e.g.* determined by applying the  
 302 CAM to a dose recovery experiment),  $\varepsilon_j$  is the analytical uncertainty as defined in Eq. (7) and  $\eta_j$  is a  
 303 centred reduced Gaussian variable as in Eq. (8). Ideally, the  $\sigma_m$  parameter should be defined by a  
 304 gamma dose recovery experiment, *i.e.* with grains all having received the same dose before loading

305 into the luminescence reader; this would presumably reproduce as closely as possible a homogeneous  
 306 irradiation in nature. However, easy access to gamma irradiations is not common in the dating  
 307 community and a beta dose recovery measurement may be the only practical alternative; in the latter  
 308 case it must be borne in mind that Thomsen *et al.* (2005) showed that such an experiment may lead to  
 309 an underestimation of  $\sigma_m$ .

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

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

$$322 \quad \Delta_j \sim \log \mathcal{N}(\mu, \sigma_d). \quad (18)$$

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

$$325 \quad \Delta = \mathbb{E}(\Delta_j) = e^{\mu + \frac{\sigma_d^2}{2}} \quad (19)$$

326 which is equivalent to

$$327 \quad \mu = \log(\Delta) - \frac{\sigma_d^2}{2}. \quad (20)$$

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

$$329 \quad \Delta_j = \Delta e^{-\frac{\sigma_d^2}{2}} e^{\sigma_d v_j} \quad (21)$$

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

$$331 \quad D_{e,j} \sim \log \mathcal{N}(\log \Delta - \frac{\sigma_d^2}{2}, \sigma_j^2 + \sigma_m^2 + \sigma_d^2) \quad (22)$$

332 which can be rewritten in the form

$$333 \quad D_{e,j} = \Delta e^{-\frac{\sigma_d^2}{2}} e^{\varepsilon_j + \sigma_m \eta_j + \sigma_d v_j} \Leftrightarrow d_j = \log \Delta - \frac{\sigma_d^2}{2} + \varepsilon_j + \sigma_m \eta_j + \sigma_d v_j. \quad (23)$$

The log-likelihood function is then given by

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

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

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

and

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

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}**

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

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

385 **{Insert Fig. 7 here}**

386 In Fig. 8, the ratio of single grain OSL to reference age is plotted as a function of reference age. It  
387 appears that the accuracy of the OSL ages decreases with age (the slope of the fitted line is  $-2.2 \pm 1.1$   
388  $10^{-3} \text{ ka}^{-1}$  for ADM-based ages and  $-2.5 \pm 1.0 10^{-3} \text{ ka}^{-1}$  for CAM-based ages and thus significantly differs,  
389 in the latter case, from zero). This trend was already observed for CAM-based ages by Guérin *et al.*  
390 (2015a; their Fig. 5b). So, while the accuracy of ADM-based ages is generally better than that of CAM-  
391 based ages, a loss of accuracy seems to be associated with increasing age. Conversely, the Bayesian  
392 model BaSAR of Combès *et al.* (2015) did not show such a trend (the slope of the line obtained by  
393 Guérin *et al.*, 2015a, is  $0.2 \pm 1.1 10^{-3} \text{ ka}^{-1}$  - cf. their discussion and Fig. 5a). This might be explained by  
394 the fact that the CAM and ADM require lists of  $D_{e,j}$  and  $\sigma_j$  values, *i.e.* simple parameterisations of  
395 individual equivalent data: the probability density of the  $D_e$  of each grain/aliquot is described by a  
396 lognormal distribution (Eqs. (9) and (16)). However, when the natural signal lies on a non-linear portion  
397 of the dose response curve (*e.g.*, close to saturation for grains having a near-zero linear component),  
398 the variance in the probability density of  $D_e$  values becomes increasingly large and the lognormal  
399 distribution may not satisfactorily describe this density. At present, we cannot think of a simple  
400 function that would describe, better than lognormal distributions, both aliquots in the linear range of  
401 the dose response curve and aliquots close to saturation. A way around this is provided, *e.g.* in the  
402 Analyst software, by the use of Monte Carlo simulations of both  $L_n/T_n$  and dose response curves  
403 (Berger, 2010; Duller, 2015). In such a case, the  $D_e$  probability density distributions are more complex;  
404 but these distributions cannot be fed into the ADM (nor into the CAM) as these are simple parametric  
405 statistical models.

406 In contrast, such complex  $D_e$  probability distributions are taken into account in the Bayesian  
407 model of Combès *et al.* (2015), which leads us to hypothesise that the BaSAR model handles larger  
408 doses better compared to the CAM and ADM. Thus, it seems that while in general quartz OSL age  
409 underestimation is a widely acknowledged concern (*e.g.* Buylaert *et al.*, 2007), at least part of it can be

410 attributed to inadequate data analysis (see also the discussion in Guérin *et al.*, 2015b). However,  
411 further testing of this hypothesis is beyond the scope of the present study.

412 *{Insert Fig. 8 here}*

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

### 414 6.1. Consequences for the use of the MAM and FMM

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

### 420 6.2. A further note on the use of the FMM

421 We now turn to the implications of this discussion on the use of the FMM. The FMM is mainly  
422 used to separate two - or more - discrete dose components in a sediment sample. These components  
423 are generally presumed to have resulted from the mixing of grains from two different layers of  
424 different ages (say, layers 1 and 2). While this assumption may be reasonable (although difficult to  
425 prove) when a stratigraphic record has been affected by post-depositional processes, the effect of such  
426 mixing has consequences on dose rates that have rarely been discussed. Deeben *et al.* (2013) noted  
427 that the dose rate experienced by the grains before mixing is unknown and may be different from  
428 today's measured dose rate (*i.e.* after mixing); as a result, the authors advocated caution in the  
429 interpretation of FMM ages. To formalise the problem, if one assumes that the mixing of layers 1 and  
430 2 occurred a time  $t_m$  ago, then we can write that:

$$431 \Delta_1 = \bar{D}_1(t_1 - t_m) + (f_1\bar{D}_1 + f_2\bar{D}_2)t_m \quad (27)$$

432 where  $t_1$  is the age of sediment deposition of layer 1, and  $f_1$  and  $f_2$  represent the proportions of layer  
433 1 and 2 in the mixing of these two layers (a similar equation can also be written for layer 2). This  
434 equation quite simply states that any mixing of sufficient magnitude to be reflected in the dose  
435 distribution must, in general, also have an impact on radioelement concentrations. We can rewrite Eq.  
436 (27) as:

$$437 t_1 = t_m + \frac{\Delta_1 - (f_1\bar{D}_1 + f_2\bar{D}_2)t_m}{\bar{D}_1} \quad (28)$$

438 since the aim of using the FMM is to determine  $t_1$  (or  $t_2$ ). In this equation, the FMM may provide a  
439 (biased, *i.e.* underestimated) value of  $\Delta_1$  and estimates of  $f_1$  and  $f_2$  (even although these estimates are  
440 likely to be significantly in error; *cf.* Roberts *et al.*, 2000; Guérin *et al.*, 2013). However, both  $t_m$ ,  $\bar{D}_1$  and  
441  $\bar{D}_2$  are unknown, (obviously one would have taken a sample from layer 1, and another from layer 2,  
442 had this been possible in the field). In other words, either the ingredients necessary for the age  
443 calculation are absent, or modelling could be avoided. As a consequence, we must regard published  
444 FMM ages as of doubtful value, except possibly in cases where dose rates do not vary significantly  
445 through the section containing the layers of interest.

448 The only exceptions to this rule are (i) if dose rates from the original Layers 1 and 2 were  
449 identical – but this assumption cannot be tested; or (ii) in the particular case where the originally upper  
450 layer was simultaneously deposited and mixed with the older layer. In such a case, the lowest dose  
451 component identified by the FMM could, in principle, be used in conjunction with the measured, mixed  
452 dose rate to calculate the age of the grains in the originally upper layer. It should be emphasised here  
453 that it is the dose value which identifies the grains that can be accurately dated by applying the FMM,  
454 rather than the proportion of grains in the various components. It is usual in the literature to select  
455 the component in which the majority of grains are found, but there is no a priori reason to expect that  
456 the measured dose rate to this component has applied throughout the burial period.

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

### 462 **6.3. Multi-grain OSL age estimates**

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

## 475 **8. Conclusion**

476 The definition of equivalent dose implies that if any overdispersion is observed in OSL dose  
477 recovery distributions, it can only result from unidentified measurement errors. Conversely, extrinsic  
478 sources of dispersion in natural  $D_e$  distributions such as those arising from dose rate heterogeneities  
479 are of a very different nature. Here we have, for the first time in the statistical analysis of single grain  
480 OSL data, discussed the effect of experimentally-measured dose rate parameters on ages. We point  
481 out that the Central Age Model (as well as the MAM and FMM) estimates the median (or geometric  
482 mean) of a lognormal distribution, whereas parameters determined for the distribution of  
483 corresponding dose rates are averages (arithmetic means). As a result, CAM-based single grain OSL  
484 doses will inevitably underestimate the burial doses and thus result in age underestimation. This model  
485 prediction has been validated by a series of experiments on samples for which independent  
486 chronological information is available. The amount of age underestimation is, for our sample set,  $8 \pm$   
487  $2 \%$ . A new model, the Average Dose Model (ADM; code available in the R ‘Luminescence’ package)  
488 has been introduced to address the identified weaknesses of the CAM. The ADM can be applied to  
489 well-bleached samples and leads to more accurate ages, which in our study are, on average, in  
490 agreement with independent ages. However, for both models an increasing age underestimation is  
491 observed with increasing age: we attribute this, at least in part, to improper analytical treatment of  
492 the effect of saturation of the quartz OSL signal with dose.

493 Finally, we argue that more appropriate acronyms should be used in the literature, such as  
494 CDM (for Central Dose Model) instead of CAM (and similarly, MDM instead of MAM). Ages calculated  
495 using the FMM should only be used with caution, and preferably avoided altogether, because the  
496 average dose rates absorbed by the grains of the different components cannot be known, except in  
497 very specific cases. For multi-grain aliquots, dividing a simple unweighted arithmetic mean of individual  
498  $D_e$  values by the average dose rate is, in general, most likely to give accurate ages.

## 500 **Acknowledgements.**

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

## 506 **References**

- 507 Aitken, M. J., 1985. Thermoluminescence dating. Academic Press, London.
- 508 Arnold, L.J., Roberts, R.G., Galbraith, R.F., DeLong, S.B., 2009. A revised burial dose estimation  
509 procedure for optical dating of young and modern-age sediments. *Quaternary Geochronology*, 4, 306-  
510 325.
- 511 Bailiff, I.K., Lewis, S.G., Drinkall, H.C., White, M.J., 2013. Luminescence dating of sediments from a  
512 Palaeolithic site associated with a solution feature on the North Downs of Kent, UK. *Quaternary*  
513 *Geochronology* 18, 135-148.
- 514 Banerjee, D., 2001. Supralinearity and sensitivity changes in optically stimulated luminescence of  
515 annealed quartz. *Radiation Measurements* 33, 47-57.
- 516 Berger, G.W., 2010. Estimating the error in equivalent dose values obtained from SAR. *Ancient TL* 28,  
517 55-66.
- 518 Combès, B., Lanos, P., Philippe, A., Mercier, N., Tribolo, C., Guérin, G., Guibert, P., Lahaye, C., 2015. A  
519 Bayesian central equivalent dose model for optically stimulated luminescence dating. *Quaternary*  
520 *Geochronology* 28, 62-70.
- 521 Cunningham, A.C., Wallinga, J. and Minderhoud, P.S.J. 2011. Expectations of scatter in equivalent-dose  
522 distributions when using multi-grain aliquots for OSL dating. *Geochronometria* 38, 424-431.
- 523 Cunningham, A.C., Wallinga, J. and Minderhoud, P.S.J. 2011. Expectations of scatter in equivalent-dose  
524 distributions when using multi-grain aliquots for OSL dating. *Geochronometria* 38, 424-431.
- 525 Deeben, J., Hiddink, H., Huisman, D.J., Müller, Schokker, J., Wallinga, J. 2010. Middle Palaeolithic  
526 artefact migration due to periglacial processes; a geological investigation into near-surface occurrence  
527 of Palaeolithic artefacts (Limburg – Eastern Brabant coversand region, the Netherlands). *Netherlands*  
528 *Journal of Geosciences – Geologie en Mijnbouw* 89, 35-50.
- 529 Dietze, M., Kreutzer, S., Burow, C., Fuchs, M. C., Fischer, M., Schmidt, C., 2016. The abanico plot:  
530 visualising chronometric data with individual standard errors. *Quaternary Geochronology* 31, 12-18.
- 531 Duller, G.A.T., 2015. The Analyst software package for luminescence data: overview and recent  
532 improvements. *Ancient TL* 33, 35-42.

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

535 Galbraith, R.F., 1988. Graphical Display of Estimates Having Differing Standard Errors. *Technometrics*  
536 30, 271-281.

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

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

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

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

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

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

552 Guérin, G., Mercier, N. and Adamiec, G, 2011. Dose-rate conversion factors: update. *Ancient TL*, 29 (1),  
553 5-8.

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

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

560 Guérin, G., Frouin, M., Talamo, S., Aldeias, V., Bruxelles, L., Chiotti, L., Dibble, H. L., Goldberg, P., Hublin,  
561 J.-J., Jain, M., Lahaye, C., Madelaine, S., Maureille, B, McPherron, S. P., Mercier, N., Murray, A. S.,  
562 Sandgathe, D., Steele, T. E., Thomsen, K. J., Turq, A., 2015c. A Multi-method Luminescence Dating of  
563 the Palaeolithic Sequence of La Ferrassie Based on New Excavations Adjacent to the La Ferrassie 1 and  
564 2 Skeletons, *Journal of Archaeological Science* 58, 147-166.

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

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

571 Huntley, D. J., Godfrey-Smith, D. I., Thewalt, M. L. W., 1985, Optical dating of sediments, *Nature*, 313,  
572 105-7.

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

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

577 Jain, M., Bøtter-Jensen, L., Thomsen, K.J., 2007. High local ionization density effects in x-ray  
578 excitations deduced from optical stimulation of trapped charge in Al<sub>2</sub>O<sub>3</sub>:C. Journal of Physics:  
579 Condensed Matter, 19, 116-201.

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

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

585

586 Mayya, Y.S., Morthekai, P., Murari, M.K., Singhvi, A.K., 2006. Towards quantifying beta  
587 microdosimetric effects in single-grain quartz dose distribution. Radiation Measurements, 41, 1032-  
588 1039.

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

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

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

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

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

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

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

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

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

609 Roberts, R.G., Galbraith, R.F., Yoshida, H., Laslett, G.M., Olley, J.M., 2000. Distinguishing dose  
610 populations in sediment mixtures: a test of single-grain optical dating procedures using mixtures of  
611 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

623 **Figure captions.**

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

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

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

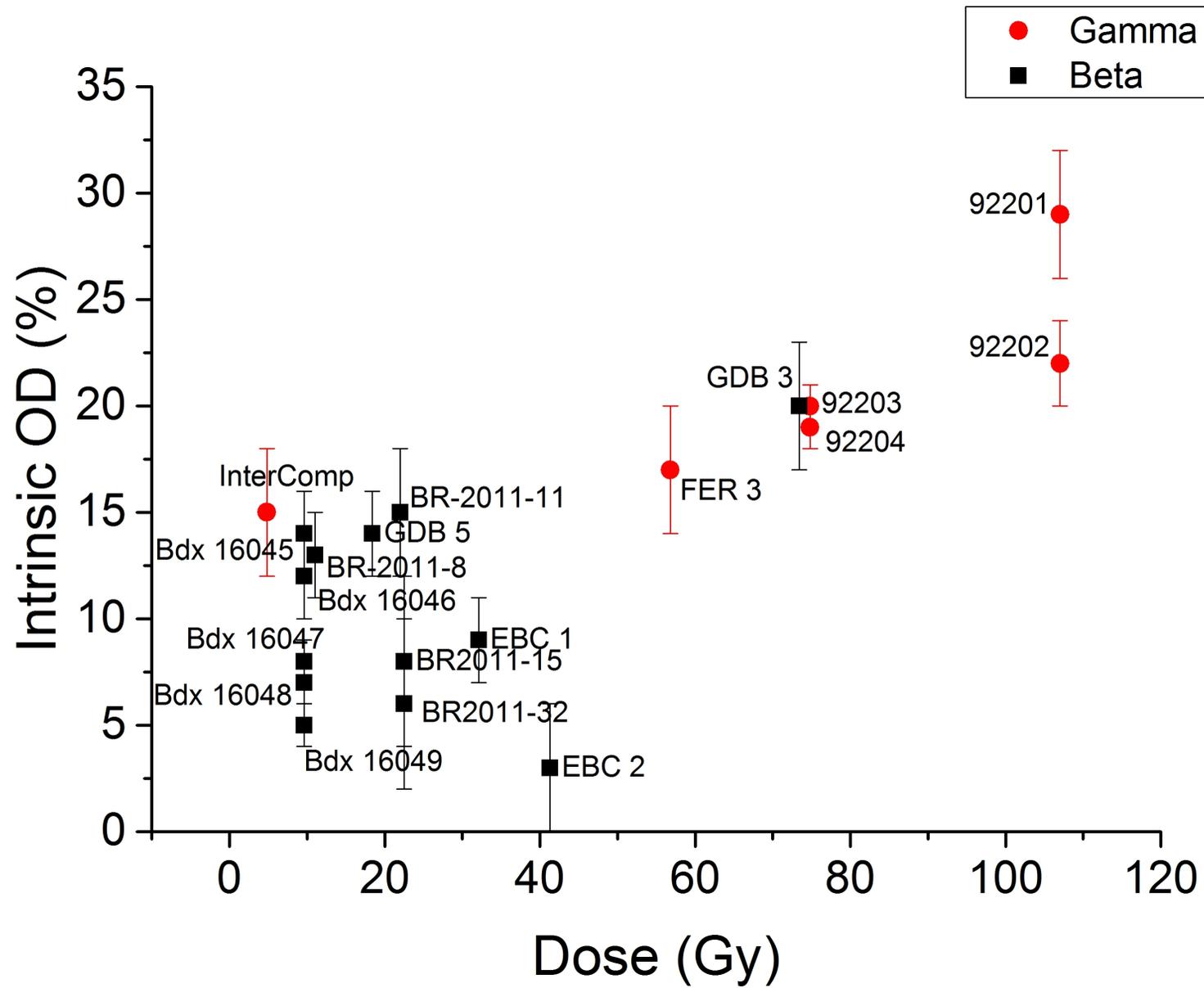
635 **Figure 4.** Radial plot showing the natural  $D_e$  distribution for sample BR-2011-8. The CAM OD has been  
636 added in quadrature to analytical uncertainties to provide the uncertainty associated with each  $D_e$   
637 estimate that is taken into account in the central dose value by the CAM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

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

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

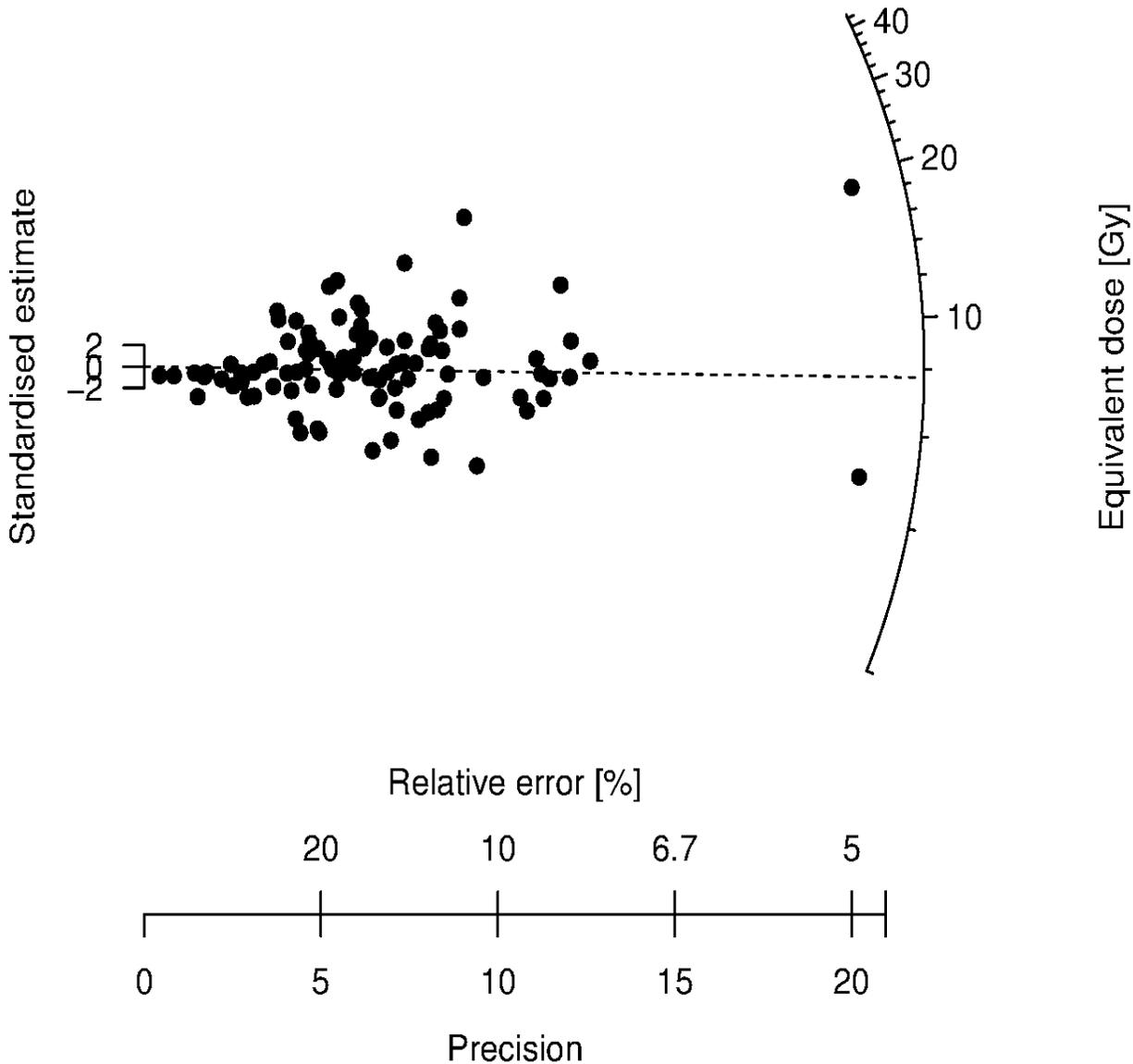
647 **Figure 7.** Relative difference between single-grain OSL CAM-based ages and reference ages, as a  
648 function of the same quantity when OSL ages are estimated with the ADM. The black line is the 1:1  
649 line: points lying above the line indicate a better performance of the ADM ( $n=14$ ); points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

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



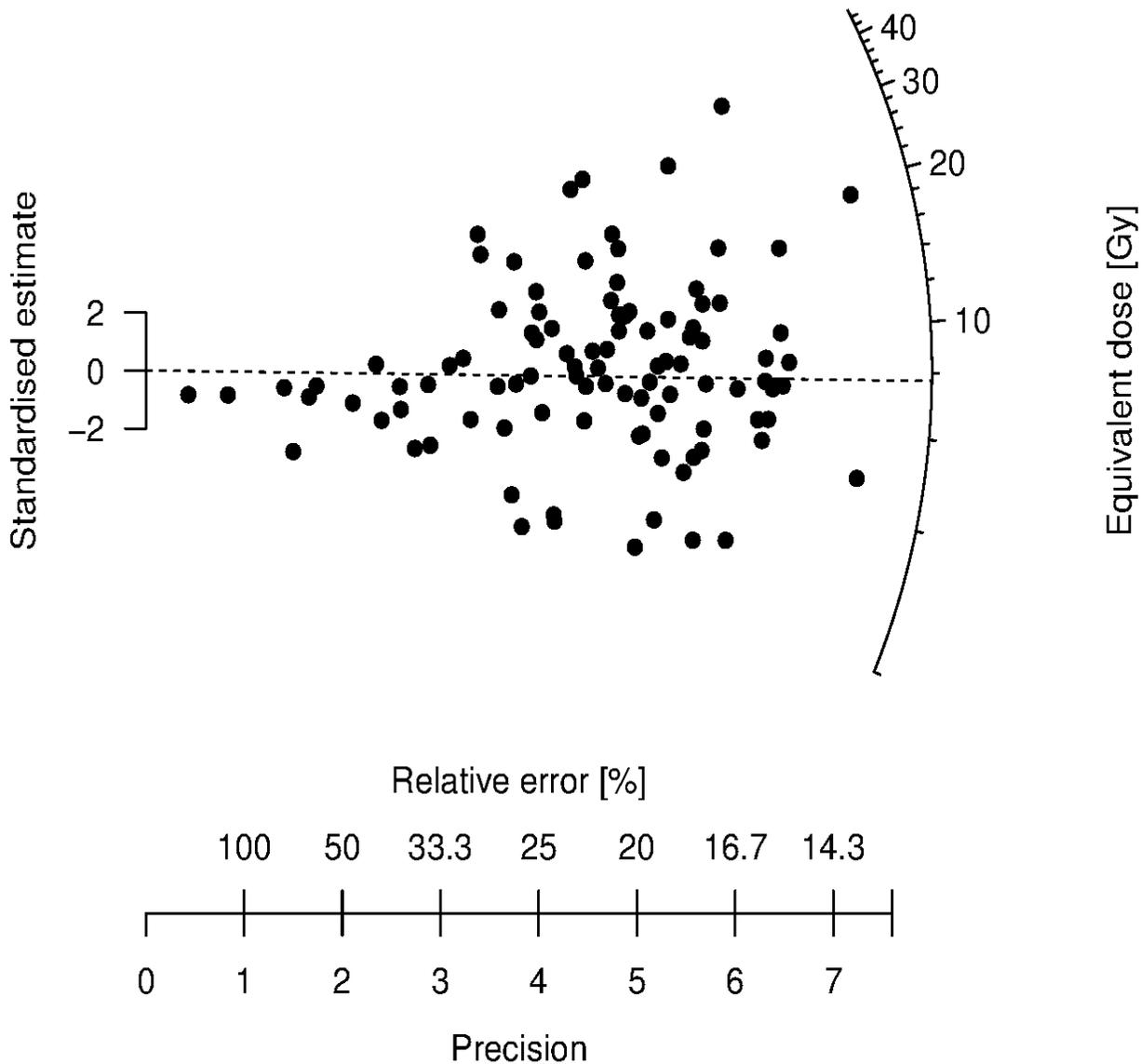
# BR-2011-8

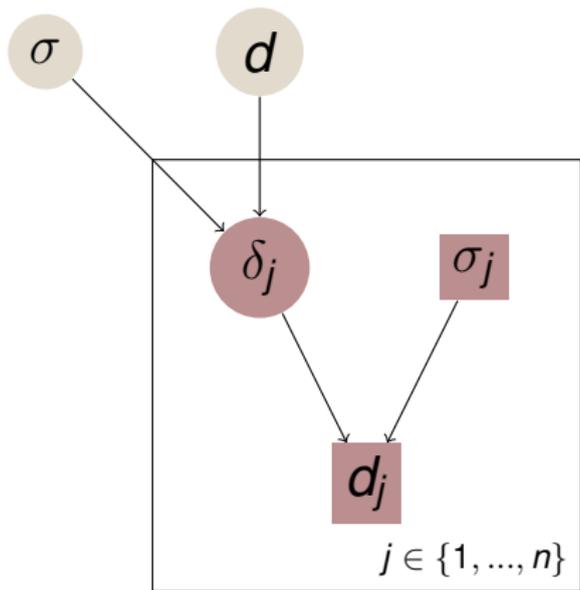
Analytical errors only



# BR-2011-8

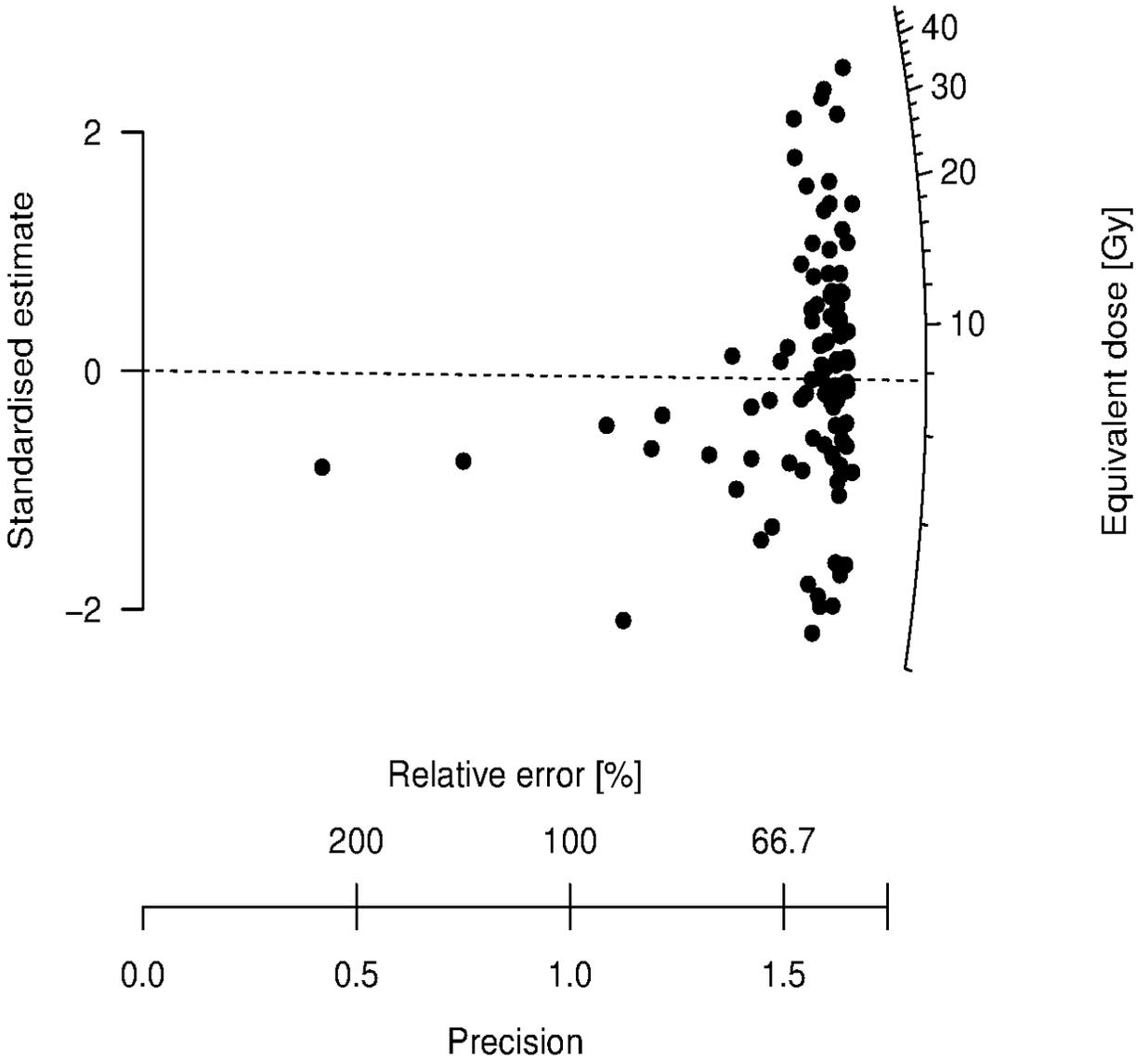
Errors including intrinsic OD

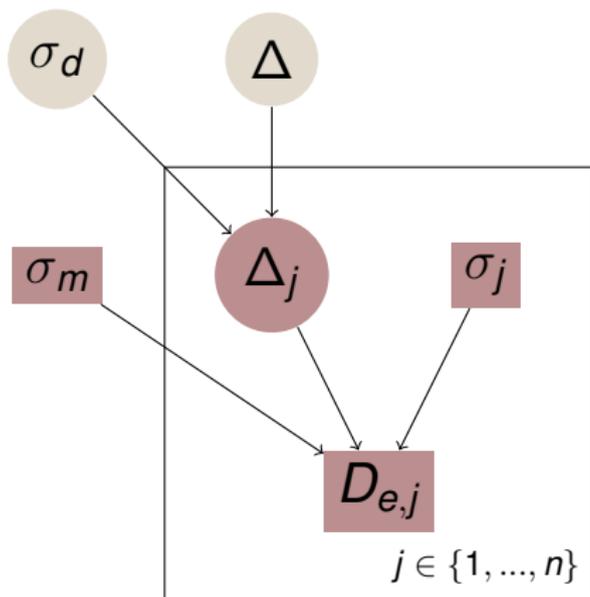


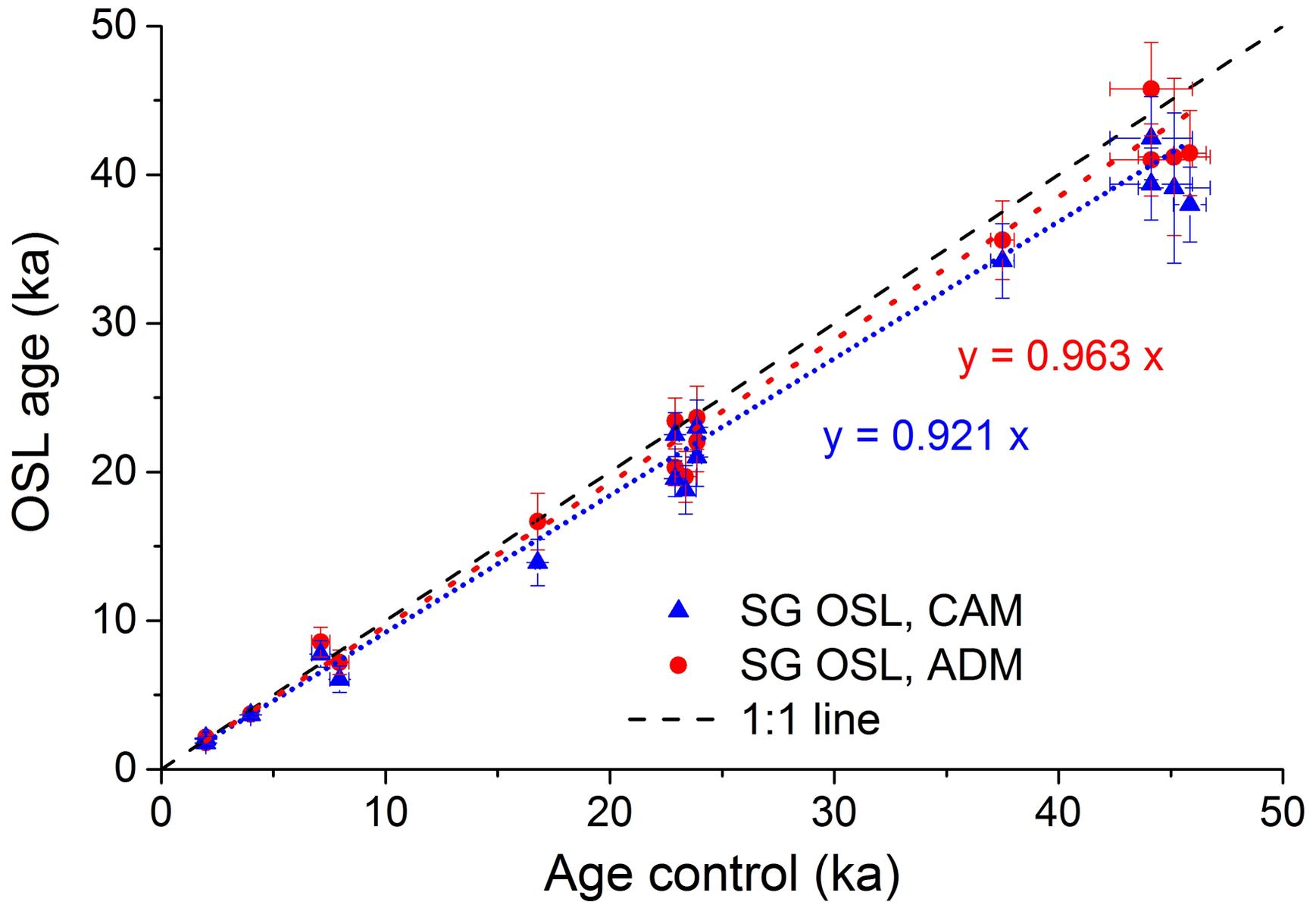


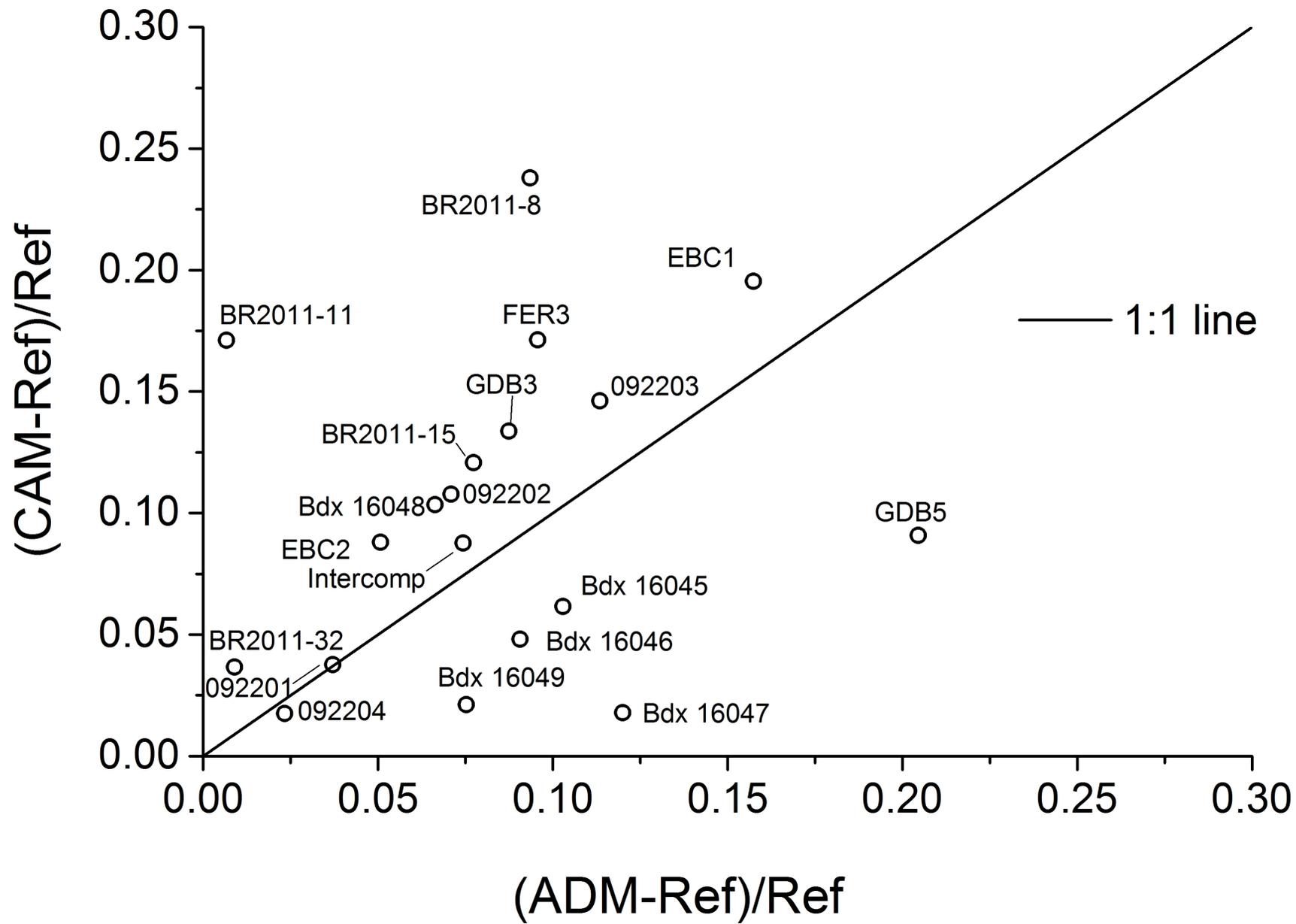
# BR-2011-8

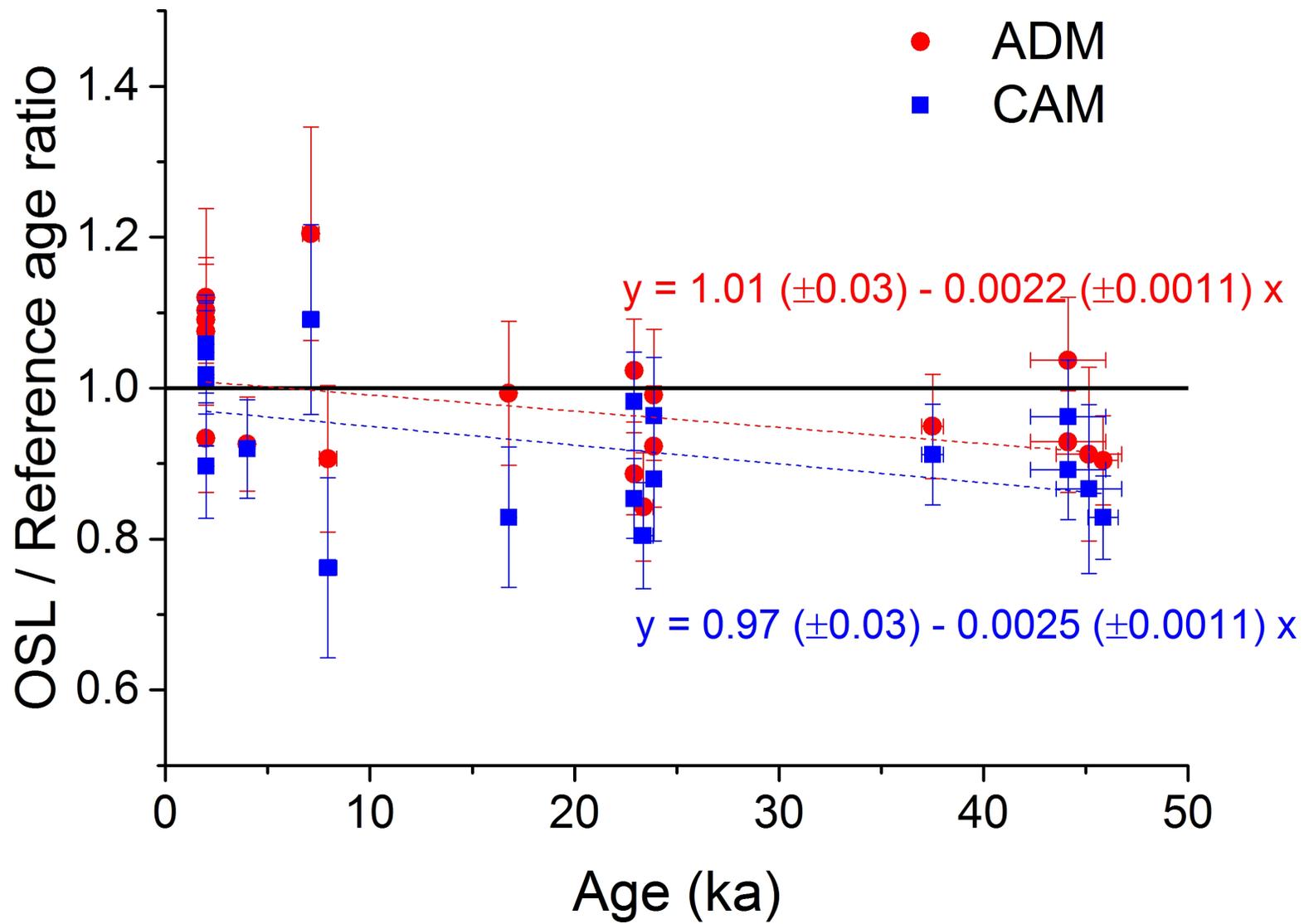
Errors including CAM OD











Sample	Given dose (Gy)	Radiation type	Meas. to given		OD <sub>int</sub> (%)	$\sigma$
			dose	$\sigma$		
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  $^{90}\text{Sr}/^{90}\text{Y}$  source in the luminescence reader. Gamma doses were delivered with a reference  $^{137}\text{Cs}$  source in Risø DTU. ‘ $\sigma$ ’ 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  $\sigma_m$ .

---

$t$ [ka]	Sample age
$\Delta_j$ [Gy]	Dose absorbed by grain $j$
$\dot{D}_j$ [Gy.ka <sup>-1</sup> ]	Dose rate to which grain $j$ was exposed
$\bar{D}$ [Gy.ka <sup>-1</sup> ]	Average dose rate, which is also the experimentally derived dose rate
$\Delta$ [Gy]	Sample palaeodose
$D_{e,j}$ [Gy]	Equivalent dose for grain $j$
$d_j$	Log ( $D_{e,j}$ )
$\sigma_j$	Relative analytical uncertainty on $D_{e,j}$
$\delta_j$	“true” logged equivalent dose for grain $j$
$D_{e,CAM}$ [Gy]	CAM dose
$d$	Log ( $D_{e,CAM}$ )
$\sigma$	CAM overdispersion
$\sigma_m$	Additional measurement uncertainty (in practice, intrinsic overdispersion)
$\sigma_d$	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	$\sigma$	Age (ka)	$\sigma$	Age ratio	$\sigma$	n	Age (ka)	$\sigma$	OD (%)	$\sigma$	Age ratio	$\sigma$	Age (ka)	$\sigma$	Age ratio	$\sigma$
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

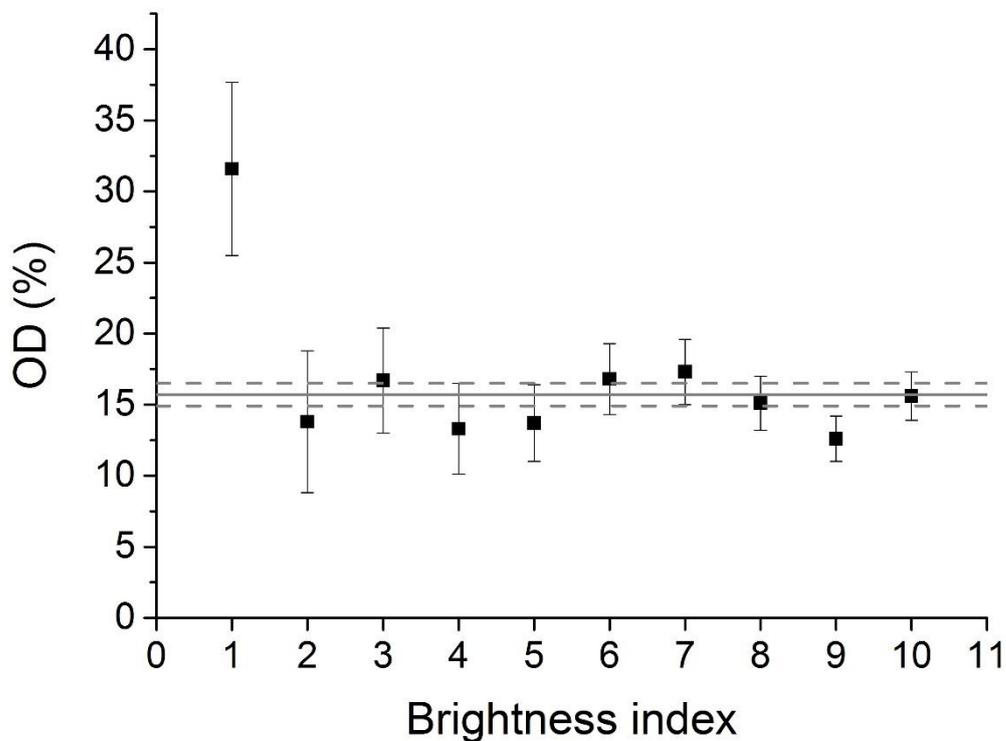
**Table 3.** Multi grain and Single grain OSL ages in comparison with independent age information (for more details see Guérin *et al.*, 2015b). ‘Ref. age’ and ‘pseudo-sigma’ stand for the age obtained independently from OSL measurements. ‘ $\sigma$ ’ denotes the standard error on the preceding column. ‘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  $^{14}\text{C}$  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.

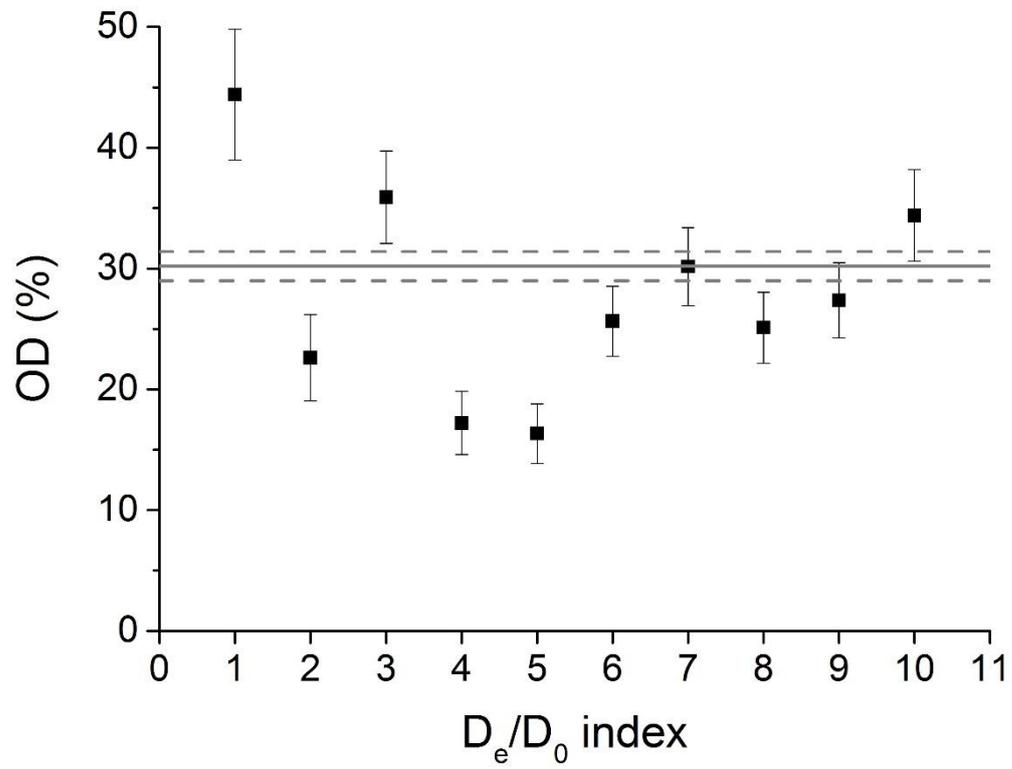
To check these assumptions, we used two datasets<sup>1</sup>: first, calibration quartz allowed us to have a large dataset with highly variable sensitivities. We sorted the grains by increasing response to the first test dose (as a proxy for sensitivity) and separated them in deciles. In Fig. 1, the 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.

<sup>1</sup> 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).