Reliability of equivalent-dose determination and age-models in the OSL dating of historical and modern palaeoflood sediments

Medialdea, Alicia; Thomsen, Kristina Jørkov; Murray, Andrew Sean; Benito, G.

Published in:
Quaternary Geochronology

Link to article, DOI:
10.1016/j.quageo.2014.01.004

Publication date:
2014

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Reliability of equivalent-dose determination and age-models in the OSL dating of historical and modern palaeoflood sediments

Keywords: Optically stimulated luminescence, quartz, flash-flood deposits, young sediments, minimum age models, single grains

Corresponding Author: Ms. Alicia Medialdea, CSIC

First Author: Alicia Medialdea

Order of Authors: Alicia Medialdea; Kristina J Thomsen; Andrew S Murray; Gerardo Benito

Abstract: The challenge of accurately estimating the deposition age of incompletely-bleached samples in luminescence dating has motivated developments in the analysis of single grain dose distributions, and a number of statistical approaches have been proposed over the last few years. In this study, we compare the behaviour of the arithmetic average, the so-called 'robust statistics', the Central Age Model (CAM), the Minimum Age Model (MAM) and the Internal-External Consistency Criterion (IEU), when applied to single-grain and small multi-grain (~30 grains per aliquot) dose distributions from a sequence of eight recent (40-1000 years old) flash-flood deposits. These sediments are expected to be incompletely bleached, but all have age control from historical records. Modifications were made to allow the use of the standard CAM and MAM models with dose distributions containing near zero and negative dose values. An assessment of minimum uncertainty on individual dose estimates is based on the over-dispersion (OD) determined in dose recovery tests making use of gamma-irradiated samples. We then present a detailed analysis of the impact of appropriate uncertainty assignment on minimum (MAM and IEU) burial dose estimates. The results of the various models are discussed in terms of the accuracy of the resulting age, and we conclude that, overall, the IEU approach generates the most accurate ages. We also demonstrate that accurate IEU ages can be obtained from multi-grain measurements if an age off-set of ~40 years can be considered to be unimportant for the samples in question. From our study we conclude that these and similar young slack-water flood deposits can be accurately dated using quartz OSL, opening up the possibility of establishing time series of flood discharge in catchments for which no instrumental or historical record exists.
Reliability of equivalent-dose determination and age-models in the OSL dating of historical and modern palaeoflood sediments

Medialdea\textsuperscript{1,3}, A., Thomsen\textsuperscript{2}, K.J., Murray\textsuperscript{3}, A.S., Benito\textsuperscript{1}, G.

1. Museo Nacional de Ciencias Naturales, CSIC, Serrano 115bis, 28006 Madrid, Spain. amedialdea@mncn.csic.es
2. Center for Nuclear Technologies, Technical University of Denmark, DTU Risø Campus, Frederiksborgvej 399, 4000 Roskilde, Denmark.
3. Nordic Laboratory for Luminescence Dating, Department of Earth Sciences, Århus University, Risø, Frederiksborgvej 399, 4000 Roskilde, Denmark.

Abstract

The challenge of accurately estimating the deposition age of incompletely-bleached samples in luminescence dating has motivated developments in the analysis of single grain dose distributions, and a number of statistical approaches have been proposed over the last few years. In this study, we compare the behaviour of the arithmetic average, the so-called ‘robust statistics’, the Central Age Model (CAM), the Minimum Age Model (MAM) and the Internal-External Consistency Criterion (IEU), when applied to single-grain and small multi-grain (~30 grains per aliquot) dose distributions from a sequence of eight recent (40-1000 years old) flash-flood deposits. These sediments are expected to be incompletely bleached, but all have age control from historical records. Modifications were made to allow the use of the standard CAM and MAM models with dose distributions containing near zero and negative dose values. An assessment of minimum uncertainty on individual dose estimates is based on the over-dispersion (OD) determined in dose recovery tests.
making use of gamma-irradiated samples. We then present a detailed analysis of the impact of appropriate uncertainty assignment on minimum (MAM and IEU) burial dose estimates. The results of the various models are discussed in terms of the accuracy of the resulting age, and we conclude that, overall, the IEU approach generates the most accurate ages. We also demonstrate that accurate IEU ages can be obtained from multi-grain measurements if an age off-set of ~40 years can be considered to be unimportant for the samples in question. From our study we conclude that these and similar young slack-water flood deposits can be accurately dated using quartz OSL, opening up the possibility of establishing time series of flood discharge in catchments for which no instrumental or historical record exists.

Keywords
Optically stimulated luminescence, quartz, flash-flood deposits, young sediments, minimum age models, single grains.
1. Introduction

Over the last decade optical dating has been increasingly used for dating Holocene fluvial deposits (e.g. Murray and Olley 2002; Wallinga, 2002; Wintle and Murray, 2006; Rittenour, 2008), complementing and extending the range of other dating techniques applicable to young deposits (e.g. radiocarbon and U-series dating). Sediments deposited by palaeofloods (Kochel and Baker, 1982) are of particular interest because they can provide high resolution records of extreme hydrological events occurring over short time periods, e.g. hours to days (Benito et al., 2004; Baker, 2008). These palaeoflood sediments are composed of stratigraphic sequences of fine-grained flood deposits emplaced in slack-water environments, usually within bedrock-controlled rivers; from these deposits the frequency and magnitude (palaeodischarge) of floods that occurred during recent centuries or millennia can be reconstructed. Palaeoflood hydrology has become important because using the stratigraphic evidence of former floods can extend the documentary record of extreme flood events considerably; this has direct application to flood hazard assessments and climate change research (Benito et al., 2004). Dating sedimentary flood units and interleaving deposits is a key task in supporting the analysis of temporal flood behaviour and recurrence (Benito and O’Connor, 2012).

In general, when dating sequences of sediments by luminescence, it is assumed that the sediment was exposed to sufficient light during transport prior to deposition to reset or bleach any latent luminescence signal. In the case of flood sediments, with short transport times and the possibility of limited light exposure, the accuracy of luminescence dating can be compromised if any latent luminescence was not completely reset during or shortly before transportation. This is likely to be a particular problem in the case of sediments transported by flash-floods (hydrographs lasting only few hours) over short distances in small basins (< 500 km²); these problems can be exacerbated by high turbidity levels or if the flood occurred at night. The consequence of such limited light exposure will be a deposit
made up of a mixture of grains with different residual luminescence signals (and so apparent residual doses). These grains are said to be incompletely reset or partially bleached. The impact of these residual doses can be compounded by the effects of thermal transfer (Rhodes, 2000; Madsen and Murray, 2009) and both these effects can lead to overestimation of the burial dose. To address the problems related to partial bleaching several statistical minimum age models have been developed for data analysis; all are intended to determine the dose recorded by a postulated population of well-bleached grains contained within a population of incompletely bleached grains.

The aim of this study is to test the accuracy of OSL dating using poorly-bleached historical and modern palaeoflood sediments. Independent age control of the individual flood layers is provided by a complete documentary record of the most extreme floods in combination with a single radiocarbon age for the oldest unit measured using OSL. This data set has been used to test the ability of various statistical models to obtain accurate burial ages for young palaeoflood sediments. Specifically we apply descriptive and robust statistics, the Central Age Model (CAM, Galbraith et al., 1999), the Minimum Age Model (MAM, Galbraith, 2005), the unlogged Minimum Age Model (MAMUL, Arnold et al., 2009) and the Internal-External Consistency Criteria (IEU, Thomsen et al., 2003; 2007) to both single-grain and small multi-grain (~30 grains) aliquot dose distributions. Single grain analysis is generally much more time consuming than multi-grain analysis (Rhodes, 2007) and it is therefore worthwhile testing whether single grain results are in fact superior to multi-grain results.

2. Site description and age framework

Sediment samples were taken from flood-related fluvial facies from the upper Guadalentin River (upstream catchment area 372 km²) in south-eastern Spain (Figure 1). The sediment package consists of sand and silt flood sediments that accumulated
from suspension during high stage floods in a slack-water environment. These slack-
water flood sediments accumulated from many successive palaeofloods and now
form an up to 7 m thick flood bench on the right margin of the Guadalentín River,
150 m upstream of a bedrock gorge entrance. At the Estrecho site, the stratigraphic
section contains a very well-developed sequence of multiple fine-grained flood
deposits, providing evidence of at least 24 individual flood layers deposited over the
last 1000 years (Benito et al., 2010, Figure 1). Five of these have been radiocarbon
dated using charcoal. Pretreatment of the sample material for radiocarbon dating
was carried out by the $^{14}$C laboratory of the Department of Geography at the
University of Zurich (GIUZ). The AMS (accelerator mass spectrometry)
measurement was made at the Institute of Particle Physics at the Swiss Federal
Institute of Technology, Zurich (ETH). Calibration of the radiocarbon dates was
carried out using the CalibeETH 1.5b (1991) program of the Institute for Intermediate
Energy Physics ETH Zürich, Switzerland, using the calibration curves of Kromer and
Becker (1993), Linnick et al. (1986) and Stuiver and Pearson (1993). The resulting ages
are shown in Figure 1.

Unfortunately the resolution of the radiocarbon calibration curve over the last 200-
300 years prevents an accurate age determination of younger material, and so only
the radiocarbon age for our oldest sample (T-17 corresponding to unit 20), lacking
historical records, was used for comparison with OSL results (although it should be
noted that a Bayesian age-depth model has been developed using the Oxcal
radiocarbon calibration software for the $^{14}$C chronology of this site, Thorndycraft et
al., 2012). These deposits contain a record of all floods exceeding the elevation of the
flood bench surface (main extreme floods) and so it is confidently expected that this
sedimentary sequence should correlate well with the historically documented flood
record in the village of Lorca, 30 km downstream. Although channel bed incision or
aggradation cannot be completely ruled out in the reach from which samples were
collected, the hydraulics and erosion base level of the channel thalweg are controlled
by the bedrock gorge located downstream. Flow accelerates (velocity over 3 m/s) in the bedrock gorge (25 m wide by 50 m high) and most of the deposits transported as bed load are flushed downstream. There is no field evidence of aggradation over the period of time covered in this paper, and the historical aerial photos since 1947 indicate a similar base level. We conclude that any channel bed changes in our study reach must have been minor (± 1 m) and should not affect the general assumption of this paper, i.e. that floods recorded in the stratigraphy correspond to extreme floods.

Based on these assumptions, Benito et al. (2010) assigned a flood year to each palaeoflood unit recorded in the stratigraphy (Figure 1) on the basis of the 500-year continuous record of extreme flooding (extracted from the Municipal Historical Archive of Lorca); this record contains information on the flood date and duration, casualties, inundated areas and damage to buildings, bridges and mills and to agricultural land, with detailed reports on the income and expenditure of the community on flood damages (Benito et al., 2010; Machado et al., 2011). Based on the recorded socio-economic damage and the geomorphological implications, floods were classified into three categories (Barriendos and Coeur, 2004): (i) Ordinary floods - restricted to the channel section (discharge lower than the bankfull discharge). These do not reach overbank sites (damage only on those items located in the channel, e.g. mills, etc). (ii) Extraordinary floods - overflow onto floodplain areas but water depth and velocity are not enough to produce widespread destruction. (iii) Catastrophic floods – these produce overbank flow with high velocities, and considerable destruction of permanent infrastructure. Thus from the qualitative descriptions in the historical record, some flood magnitude component can be extracted. It is also likely that only floods classified as (ii) or (iii) are expected to be recorded in the flood stratigraphy, but we must also consider possible sources of error in this assumption.

Although sediment yield is important in relation to the thickness of the palaeoflood units (not discussed in detail in this paper), it plays only a secondary role in the
presence or otherwise of slack-water flood deposits. In contrast, flood stage (discharge) plays a major role in the presence of a unit, because only floods exceeding the elevation of the flood bench can lead to the deposition of a flood unit (Benito and Thorndycraft, 2005). So the floods documented in the historical record are likely to have resulted in a sedimentary unit, but is that unit likely to have been preserved?

Benito et al. (2010) point out that the depositional site and the textural characteristics of the slack-water flood deposits indicate that the material was deposited in an eddy circulation during flood events. The breaks between flood units are horizontal and no erosive contacts have been identified in the sequence. Most of the units finish up with a fine laminae of silt or clay; these layers have usually been later bioturbated but their preservation indicates that the depositional environment is not prone to erosion even during major events. We conclude that it is very likely that (i) the historically recorded floods resulted in deposition, and (ii) that these depositional units had a high preservation potential.

During the period AD1500-1900, 31 flood events were reported in the Lorca Municipal Archive, of which 15 were classified as ordinary, 7 extraordinary and 9 catastrophic (Benito et al., 2010). During the second half of the 19th century, an increase in frequency of extraordinary and catastrophic floods is reflected both in the documentary record (with high economic losses and casualties) and in the palaeoflood record (with high energy sedimentary structures). As discussed in Benito et al. (2010), the stratigraphic record at the Estrecho site contains flood layers with high-energy sedimentary structures (in units 17, 18, 21 and 22) matching the most extreme recent floods documented in the Lorca archive. The most severe historical floods occurred in 1838 and 1879 and are assigned to units 18 and 21, respectively. Between these floods, the stratigraphy shows at least two flood units (units 19 and 20; the former < 7 cm in thickness) whereas the documentary records contains only one catastrophic flood (1860); Benito et al. (2010) assigned an additional
documented flood in 1877 (classified as ordinary from the Lorca record) to unit 20 to account for this extra event. The largest flood at the end of the 19th Century occurred in 1891, which was assigned to unit 22. The second largest flood in the 20th Century occurred in 1941 (palaeoflood unit 23) and the largest in 1973 (unit 24). These ages, assigned to each sedimentary flood unit on the bases of documentary written sources, are considered to provide good independent age control against which the robustness of OSL statistical age models can be tested.

3. Methodology

3.1. Sample preparation and dose rate determination

A total of eight stratigraphic units from the Estrecho site were sampled for OSL dating (Figure 1). Samples were wet sieved to grain size fractions 90-180 µm and 180-250 µm and chemically treated to isolate the quartz fraction. Samples were treated with 10% HCl until carbonates were dissolved, and then with H₂O₂ to remove organic matter. Density separation was not used on these samples; a solution of 40% HF was used to dissolve all feldspar and to remove the outer quartz layer affected by alpha radiation. The samples were resieved after chemical treatment to give grains in the range 180-250 µm. Feldspar contamination was tested for using IR simulation prior to OSL measurements; no detectable IRSL signal was observed.

Dose rates are based on the average radionuclide activities of material collected from the hole remaining after removal of the sampling tube. High resolution gamma spectrometry was used to measure concentrations of ²³⁸U, ²³⁵Th and ⁴⁰K (Murray et al. 1987). Appropriate conversion factors (Olley et al., 1996) were then used to derive the dose rates. The site is well-drained and all deposits are well above the level of normal flow. The elevation of the river channel is related to the local base level; this is controlled by the bedrock gorge located 500 m downstream from the stratigraphic
profile. As discussed above, only minor changes in the river level (~1 m) are considered likely.

It is assumed that the water content at the time of sampling is representative of the average site water content. For simplicity a linear accumulation of deposits has been assumed in order to calculate the contribution of cosmic radiation according to a varying burial depth (based on Prescott & Hutton, 1994). Sampling depths below modern surface, observed water contents, radionuclide activity concentrations and derived dry total dose rates to an infinite matrix are summarized in Table 1.

By combining the dose rate data with independent age control (see section 2) the expected burial doses for each unit have been calculated from these known ages and the corresponding expected doses are also summarized in Table 1.

3.2. Instrumentation

Quartz OSL measurements were made using automated Risø TL/OSL DA-20 luminescence readers (Bøtter-Jensen et al., 2010). Luminescence was detected using a bialkali EMI 9235QB photomultiplier tube through a 7.5 mm Hoya U-340 filter. Stimulation of multi-grain aliquots used a blue (470 ± 30 nm) light emitting diode array providing a stimulation power of ~80 mW/cm² at the sample position. Grains were mounted as a monolayer on 9.7 mm diameter stainless steel discs using silicone oil. Single grain measurements used 180-250 µm diameter grains, and were undertaken using the Risø single grain attachment (Duller et al., 1999; Bøtter-Jensen et al., 2000). Single grains were loaded into aluminium discs containing 100 grain holes, each with a depth and diameter of 300 µm. Optical stimulation used a 10 mW green (532 nm) Nd:YVO₄ laser providing a power density of ~50 W/cm² at the stimulation position.
Laboratory irradiations were undertaken using either in situ calibrated $^{90}$Sr/$^{90}$Y beta sources (providing dose rates to quartz mounted on stainless steel discs of 0.04 to 0.11 Gy/s) or a $^{137}$Cs (662 keV) collimated point gamma source in a scatter-free geometry (providing an accurately known dose rate in quartz of ~ 30 μGy/s). Quartz grains were gamma irradiated in the dark in glass tubes (ID = 2.6 mm and wall thickness = 1.8 mm).

3.3. OSL analysis method

The quartz 180-250 μm fraction was chosen for all measurements as larger grain sizes are expected to be better bleached than smaller grain sizes (Olley et al., 1998; Colls et al., 2001; Truelsen et al., 2003; Fuchs et al., 2005) despite being expected to be exposed to less light during transport (Wallinga, 2002). This is also the preferred grain size for single grain measurements. Samples were measured using the SAR procedure (Murray and Wintle, 2000; 2003) with three to five regeneration doses, a zero dose point to check recuperation of the signal after the highest regeneration dose measurement cycle, and a recycling of the second regeneration dose. Our protocol employed a preheat temperature of 200°C for 10 s and a cutheat of 180°C at a heating rate of 5°C/s for all measurements (see section 4.3).

Multi-grain measurements were carried out using 2 mm aliquots containing ~30 grains each. The number of grains on each aliquot was assessed under a microscope. The OSL signals from multi-grain aliquots are based on the summation of the first 0.64 s of stimulation corrected for background derived from the following 0.64 s in order to eliminate the contribution of other components of smaller optical cross-sections (Early BackGround subtraction, EBG, e.g. Cunningham and Wallinga, 2010). Single-grain OSL signals are derived from the summation of the first 0.1 s of stimulation less the sum of the last 0.2 s (i.e. Late BackGround subtraction, LBG).
Using EBG on single grain measurements reduced the number of accepted grains by 
~30% without changing the dose estimate or over-dispersion significantly (see 
section 5.1.1 for further details) – this may reflect variable effective stimulation 
power arising from reflection and scatter from grain surfaces (Thomsen et al., 2012). 
Poisson statistics were assumed when estimating the uncertainty on the background-
corrected signal.

Dose estimates were accepted (both multi-grain and single grain measurements) if 
the relative uncertainty on the natural test-dose response was less than 20% and if 
the recycling value, excluding uncertainties of individual grains, was within 20% of 
unity. Aliquots having a dose response curve with anomalous behaviour (e.g. 
decreasing response for increasing dose) were also discarded (~5% of the grains for 
otherwise acceptable single grains, <1% for multi-grain aliquots). These criteria led to 
the rejection of 98% of the measured single grains and ~5% of the multi-grain 
aliquots. Doses and their uncertainties were estimated using Analyst 3.24 (Duller, 
2007). Equivalent doses ($D_e$) were estimated by interpolation of the natural test-dose-
corrected signal onto the dose response curve, fitted using either a linear or a 
saturating exponential function (see insets in Figure 2). Approximately 2% of the 
measured sensitivity-corrected natural signals were higher than the signal measured 
for the highest regeneration dose (70 Gy); in these cases the $D_e$ was determined by 
extrapolation. The uncertainties derived from Analyst 3.24 are based on counting 
statistics and curve fitting errors. An additional uncertainty derived from gamma 
dose recovery experiments was added onto individual dose estimates to account for 
additional measurable intrinsic sources of variability (see section 4.4, Thomsen et al., 2005, 2007, 2012; Reimann et al., 2012; Sim et al., 2012).

All single grain dose distributions, except those obtained after in situ beta 
irradiation, were corrected for beta source inhomogeneity using the approach 
described by Lapp et al. (2012, see section 4.5).
3.4. Burial dose estimation

To estimate the burial dose we use five different approaches:

(i) The simple unweighted arithmetic mean and standard deviation. In this case all dose estimates are included in the burial dose estimation.

(ii) The unweighted arithmetic mean, but with data selected using robust statistics (Tukey, 1977). In this approach extreme outliers are removed prior to calculation of the unweighted mean. Extreme outliers are identified to be those outside the $1.5 \times \text{IQR}$ (InterQuartile Range), where $\text{IQR} = Q_3 - Q_1$ (first quartile $Q_1$ separates the lowest 25% of the data, third quartile $Q_3$ separates the highest 25% of the data, Tukey, 1977). Unweighted averages and the corresponding standard deviations are then calculated for the resulting distributions.

(iii) The Central Age Model, (CAM, Galbraith et al., 1999) which calculates the weighted average including all doses estimates. Most of the dose distributions presented here contains some non-positive dose estimates which prevent the direct use of CAM because of its log-normal assumption. Arnold et al. (2009) described a version of the CAM, CAM$_{UL}$, in which the log-normal assumption has been replaced by a normal assumption and thus is able to deal with dose distributions containing non-positive dose estimates. However, by using a simple exponential transformation of the data prior to application of the standard CAM, we obtained identical results to those derived from the CAM$_{UL}$. If the individual dose values and their associated uncertainties are termed $D_i$ and $s_i$, respectively, then the transformed values are

(iv) The Minimum Age Model (MAM, Galbraith et al., 1999), which assumes that only a proportion of the measured doses belongs to the burial dose distribution and that the remaining dose estimates are part of a log-normal distribution truncated at the burial dose. As with the CAM, the MAM has an underlying log-normal assumption which prevents its application to dose distributions containing non-
positive dose estimates. Arnold et al. (2009) developed a version of MAM, MAM\textsuperscript{UL}, which is based on the assumption that the dose distributions are normal. However, applying the simple exponential transformation described above also allows the use of the original (and in our experience more stable) MAM scripts. Results derived in this way are termed MAM\textsuperscript{tr}.

(v) The Internal-External Consistency Criterion (IEU; Thomsen et al., 2003; 2007) used to identify the lowest normal-dose population; this is presumed to be the population of grains most likely to have been well-bleached at deposition.

(vi) The lowest 5\% $D_e$ values selection suggested by Olley et al., (1998); this approach is included in the decision tree process of Bailey and Arnold, 2006.

4. Luminescence characteristics

In this section we provide details of the luminescence characteristics of these quartz samples.

4.1. Decay curves and dose response curves

Figure 2a shows a normalized OSL decay curve from sample T-31 measured using multi-grain (~30 grains) aliquots; this curve is representative of those from all samples measured here. The natural OSL signal decays to half its initial value in ~0.14 s and reaches background level in ~0.8 s. Figure 2a includes a normalized OSL decay curve obtained from calibration quartz, known to be fast component dominated. A comparison of these two curves shows that the OSL signals from the samples investigated here are also fast component dominated. The average multi-grain test dose intensity is $700 \pm 120$ counts/Gy integrated over the first 0.64 s. The inset in Figure 2a shows a typical dose response curve from multi-grain aliquots, fitted with a linear function. The corresponding data based on single grain results
are shown in Figure 2b. In this case the dose response curve is fitted using a saturating exponential function but the natural signals generally fall in the linear part of the dose response curves. As is often reported in the literature, the decay shapes of the single grain OSL stimulation curves vary significantly from grain to grain (e.g. Duller, 2008 and references therein). The average test dose intensity from accepted grains is 400 ± 100 counts/Gy integrated over the first 0.1 s. No significant differences in terms of luminescence characteristics have been observed among the samples investigated here.

4.2. Cumulative light sum

The cumulative light sum (Duller and Murray, 2000) is shown in Figure 3 both for the natural signal and for the signal from the second regeneration dose (3 Gy) for sample T-26. If all grains contributed equally to the OSL, then the cumulative light sum would be directly proportional to the proportion of grains, with a slope of unity. In our samples ~80% of the total light is derived from less than 12% of the grains and ~60% of the grains do not contribute significantly to the detected signal, providing less than 10% of the total light sum. A significant proportion of grains with detectable signals is not taken into account in single grain analysis due to the selection criteria; these grains would, of course, contribute to the signal from a multi-grain aliquot.

4.3. Effect of preheat temperature

To select an appropriate thermal treatment, preheat plateau tests on multi-grain aliquots were measured on samples T-23 and T-39. Based on these measurements, it can be concluded that any preheat temperature below 300°C would be appropriate (see Figure 4a).

Thermal transfer was assessed by bleaching 48 previously unused multi-grain aliquots of each of samples T-17, T-26 and T-39 in a daylight simulator (Hönle SOL 2).
at a distance of 80 cm from the lamp for two hours prior to measurement. The results from CAM$_{tr}$ average are shown in Figure 4b; each point is the average and standard error of six individual dose estimates. For preheat temperatures less than 240°C the CAM$_{tr}$ average is consistent with zero; even at a preheat temperature of 300°C the measured dose is only ~0.25 Gy. Figure 4c shows the CAM$_{tr}$ doses from similar thermal transfer experiments measured using single grains. A total of 1500 single grains of sample T-30 were measured using preheat temperatures varying between 160 and 260°C. For this sample the initial bleach was done in the reader using two blue light exposures at room temperature for 40 s with an intervening pause of 10,000 s. The CAM$_{tr}$ doses for this sample vary between -2 ± 8 mGy to 76 ± 25 mGy for the preheat temperature range considered here. Based on these measurements and the results of the preheat plateau test we chose to employ a preheat temperature of 200°C and a cutheat temperature of 180°C; this combination is commonly used for young samples (e.g. Madsen et al., 2007; Hu et al., 2010).

Finally, 2400 grains of sample T-39 were bleached in the daylight simulator for 2 hours and measured using a preheat temperature of 200°C. Of these, 57 grains were accepted, and the resulting dose distribution (see inset to Figure 4c) appears to be symmetric around a weighted mean dose of 16 ± 7 mGy. This weighted dose is consistent with that of sample T-30 measured at a preheat temperature of 200°C (12 ± 10 mGy, n=27). Combining the results from all the bleached single grains measured at a preheat temperature of 200°C gives a weighted average of 15 ± 6 mGy (n = 84); this corresponds to a residual age of 16 ± 6 years (using the average dose rates for these two samples). This value is small, and may not be significantly different from zero. We conclude that thermal transfer is negligible in these samples when using our chosen protocol.

4.4. **Dose recovery**
As a final test of the luminescence characteristics and our measurement protocol, we have carried out dose recovery tests on both multi-grain aliquots (samples T-17, T-30 and T-39) and single grains (sample T-39) using a preheat temperature of 200°C and a cutheat of 180°C to investigate whether a known dose given prior to any thermal treatment can be recovered accurately. In these experiments all samples were bleached for two hours in the daylight simulator prior to dosing.

In one set of experiments, seven bleached multi-grain aliquots from each of the three samples were given a beta dose (in the reader) of 2 Gy. The average dose recovery ratios obtained were 0.91 ± 0.07, 0.99 ± 0.04 and 0.90 ± 0.02 (n=7) for samples T-17, T-30 and T-39, respectively. Combining the results from the three samples gives an average weighted dose recovery ratio of 0.93 ± 0.03 (n=21) and an over-dispersion (OD) of 4 ± 4 %, indistinguishable from zero.

In a different dose recovery experiment, a portion of the bleached sample T-39 was given a gamma dose of 2 Gy prior to multi-grain measurement. The dose recovery ratio for this sample was 1.07 ± 0.03 (n=48) with an OD of 7 ± 4 %. The over-dispersion derived from this experiment is consistent with that derived from the beta dose recovery experiment.

In a third experiment, 1,000 single grains of the bleached portion of sample T-39 were measured after being given a beta dose of 2 Gy in the reader. The dose recovery ratio was 0.98 ± 0.09 (n=18) with an OD of 16 ± 8 %.

Finally, 3200 grains of the gamma irradiated portion of sample T-39 gave a weighted dose recovery ratio of 0.98 ± 0.04 (n=66) with an OD value of 22 ± 5 % (after taking into account beta source inhomogeneity, see below).

From the gamma dose recovery experiments it would appear that assigning uncertainties based on counting statistics and curve fitting errors alone is insufficient to describe the observed variability in an sample known to have been uniformly irradiated. A similar conclusion was reached previously by Thomsen et al. (2005;
For all multi-grain and single grain measurements we add in quadrature an additional uncertainty (additional to that calculated from counting statistics and curve fitting errors) of 7% and 22%, respectively. This additional uncertainty is added to account for the variability in measured dose distributions arising from additional intrinsic factors; we regard this as the minimum uncertainties with which any dose can be measured using a single aliquot or single grain. If extrinsic sources of variability (e.g. small scale beta dose rate heterogeneity) contribute significantly to the spread in the well-bleached part of the dose distributions, our additional intrinsic uncertainty (7% for small aliquots and 22% for single grains) is likely to underestimate the actual variability in a well-bleached dose distribution.

### 4.5. Effect of beta source inhomogeneity

It has been reported (Spooner et al., 2000; Thomsen et al., 2005; Ballarini et al., 2006; Lapp et al., 2012) that the laboratory dose rate from the in-built beta source can be inhomogeneous across the sample area due to the position of the radioactive material on the active surface of the beta source. This effect is usually not important when measuring multi-grain aliquots but can contribute significantly to the observed variability in single grain dose distributions.

Two different beta sources were used for the single grain measurements. One of them, source ID155, was manufactured before 2000 and is thus expected to have a relatively uniform distribution of radioactive material while ID195 was manufactured in 2006 and is thus expected to have a wider spatial variation (Lapp et al., 2012). Mapping the beta source following the approach of Lapp et al. (2012) showed that the effective dose rate varied spatially by up to 5% and 13% for the sources ID155 and ID195, respectively. Lapp et al. (2012) have developed software that makes it possible to correct for this dose rate inhomogeneity automatically using this dose-rate map. This source inhomogeneity contributes to the intrinsic over-
dispersion discussed in the previous section; the size of this contribution can be
determined by analysing a bleached and gamma dosed sample with and without
correction for inhomogeneity. The observed OD is $29 \pm 6\%$ without correction for
inhomogeneity and this decreases to $22 \pm 5\%$ when the correction is employed,
suggesting that beta source homogeneity contributes $19 \pm 8\%$ to the total observed
OD of 29%.

5. Natural dose distributions

5.1. Single grain results

Between $3,000$ and $5,000$ grains from each sample were measured, resulting in 80-100
dose estimates per sample passing the rejection criteria described in section 3.3.

Natural dose distributions showing the OSL response from the natural test dose as a
function of the measured dose are shown in Figure 5. All dose distributions appear
to be positively skewed with a minimum dose edge possibly indicating the presence
of incomplete bleaching. Simple visual inspection of the dose distributions suggests
that more than 60% of all grains may belong to single dose distributions centred at
the low end of the observed values.

5.1.1. Effect of the chosen background summation limits

As discussed in section 3.3, a late background (LBG) subtraction has been used to
obtain the net OSL signal from single grains; this is in contrast to the use of an early
background, EBG, usually expected to minimise any effects of slower components
on the net signal (Cunningham and Wallinga, 2010). A comparison of the results
using EBG and LBG has been undertaken using single grains from sample T-31.
With EBG, 2.2% of the measured grains were accepted, giving a CAM$_{tr}$ burial dose
and over-dispersion of $1.5 \pm 0.6$ Gy and $300 \pm 50\%$. Estimating the minimum burial
dose using IEU gives $0.152 \pm 0.014$ Gy ($n = 49$). The use of LBG results in the
acceptance of 3.1% of the grains, giving a CAM$_{tr}$ burial dose of $1.6 \pm 0.5$ Gy and over-
dispersion of 280 ± 30 %. Using IEU on this dose distribution gives a minimum burial dose of 0.134 ± 0.012 Gy (n = 54). Thus there does not appear to be any significant difference between the estimated burial doses using either EBG or LBG for single grains (consistent with the results reported by Reimann et al., 2012). The over-dispersion of the dose distributions obtained using both EBG and LBG has been calculated for all eight samples. Again no significant differences in the ODs are observed for any one sample, although there is a slight systematic tendency for lower OD values when using LBG. More importantly, EBG results in a significant reduction (~30%) in the number of accepted grains. Thus for these samples, it would appear that using LBG gives rise to a larger data set with no detectable cost in data quality.

5.1.2. Single-grain burial dose estimates

As a first approach to burial dose estimation we calculate a simple unweighted average dose including all dose points. The results are given in Table 2 and shown in Figure 6a. Such a simple approach leads to an average over-estimation of the expected dose of ~1 Gy (n = 8), consistent with the range of values given in many reports of residual doses in modern water-lain deposits (see recent summary in Murray et al., 2012).

In an attempt to improve this simple average, we applied “robust statistics” where extreme outliers are eliminated following the 1.5*IQR criterion described in section 3.4; because of the shape of the dose distributions of these samples, all outliers identified following this criterion belong to the higher dose region of the distributions; no low dose values were rejected. The average overestimate compared to expected doses is reduced from ~1 Gy to ~0.15 Gy. Samples T-36, T-23 and T-17 now give results consistent with the expected doses within two standard deviations. The remaining samples overestimate (~0.3 Gy) the expected doses, with the largest overestimate occurring for the two youngest samples (T-39 and T-37). Thus,
although the interquartile range criterion is not based on any physical process
model, it provides a significantly better-constrained estimate of the upper dose limit
to the burial dose estimate than the simple average.

The next step was to derive weighted mean equivalent doses using the CAM model
on the exponentially transformed data, \( \text{CAM}_{tr} \). Again the doses overestimate the
expected values, by \( -0.9 \) Gy on average. Clearly, the dose overestimates do not arise
simply because of large, poorly known dose values.

As these samples are expected to be affected by incomplete bleaching because of the
nature of the deposition process and the shape of the measured dose distribution, we
expected to have to apply minimum age models to determine the true burial dose. In
the following we apply the MAM model using exponentially transformed data,
\( \text{MAM}_{tr} \), (see section 3.4) because of the presence of non-positive dose estimates in all
our dose distributions (except that from sample T-17, the oldest sample). Prior to the
application of the MAM\(_{tr}\) model an additional uncertainty of 22\% has been added to
the uncertainty of all dose estimates calculated from counting statistics and curve
fitting errors (see section 4.4) to account for intrinsic sources of uncertainty including
instrument reproducibility. This additional uncertainty of 22\% was added as a
percentage of the individual dose estimates. For both minimum age models, MAM\(_{tr}\)
and IEU, all data sets have been corrected for beta source inhomogeneity. The
minimum doses are given in Table 2 and shown in Figure 6a. Applying the MAM\(_{tr}\) to
the dose distributions we obtain agreement with the expected doses (within two
standard errors) for four of the samples (T-36, T-26, T-23 and T-17) which have
measured to expected dose ratios of \( 1.56 \pm 0.29, 1.28 \pm 0.15, 1.08 \pm 0.10 \) and \( 1.00 \pm 0.11 \).
There is a clear trend of increasing overestimation as a function of decreasing
expected dose - the best agreement is obtained for samples older than \( \sim 350 \) years.
The MAM\(_{tr}\) burial dose estimates for the remaining samples are in poor agreement
with the independent age control.
To test the IEU approach, we first fitted a straight line between the absolute OD as a function of the estimated dose obtained for the 2 Gy gamma dose recovery experiment and that obtained for the distribution of the thermal transfer experiment (0 Gy dose, see Thomsen et al., 2007). The relation obtained for OD as a function of dose is given by $OD = 0.2073 \cdot D + 0.0183$, where $D$ is the given dose. This OD was added in quadrature to all single grain data before iteratively applying the IEU model as described by Thomsen et al. (2007) to determine the IEU burial dose. The resulting IEU burial doses are in good agreement with all predicted doses; the estimates are within one standard error for four of the eight samples (T-37, T-36, T-31 and T-17) and within two standard errors for the other four.

Finally we investigated how sensitive our estimated MAM$_{tr}$ and IEU burial doses for samples T-39 and T-31 are to the size of the assigned additional uncertainty which we varied from 5% to 50%. The results are shown in Figure 7a and 7c, where it can be seen that the MAM$_{tr}$ overestimates the expected doses even for the lowest additional intrinsic uncertainty tested (5%); we are, of course, confident that the intrinsic OD must be considerably greater than 5% as the instrument reproducibility alone has been assessed to be ~6% (e.g. Thomsen et al., 2005). In contrast, the IEU results are relatively insensitive to added uncertainty, giving doses within two standard deviations of the expected values for additional uncertainties in the range 5 to 40% for single grain measurements.

Burial doses estimated using the different approaches are summarized in Table 2, together with the number of grains, $n$, included in each case. For robust statistics $n$ is the number of grains remaining after removing outliers. For MAM$_{tr}$ and IEU, $n$ is the number of grains identified as belonging to the well bleached part of each dose distribution. The estimated doses plotted against the expected values are shown in Figure 6a.

5.2. Multi-grain results
Between 80 and 100 multi-grain aliquots (~30 grains each) were measured for each sample. Laboratory gamma dose recovery tests gave an OD of $7 \pm 4\%$ ($n = 48$; see section 4.4) and so we add an additional uncertainty of 7% to individual dose estimates to account for intrinsic sources of variability. Observed natural dose distributions and the corresponding expected doses are shown in Figure 8. All distributions appear to be positively skewed and include a number of dose values significantly higher than those expected. Sample T-37 is unusual, in that even the leading dose edge seems to be off-set to higher doses compared to the expected dose. This observation is consistent with the single grain IEU analysis (see Table 2) which concluded that T-37 was the most poorly bleached sample with only ~20% of the grains identified to be well-bleached (see section 5.1.2).

The over-dispersion values for the natural samples are very similar to those determined for the single grain dose distributions (see Table 2). The simple averages, as well as the CAM$\tau$, overestimate the expected dose by ~3 Gy on average. Although it can usually be argued that it is inappropriate to apply statistical age models to multi-grain data sets, in this case, given the conclusion from the cumulative light sum, we analyse the multi-grain aliquot data set in the same manner as the single grain data.

The ratio between the MAM$\tau$ multi-grain burial dose and the expected dose is only consistent with unity (within two standard errors) for the oldest sample (T-17), which has a ratio of $0.93 \pm 0.26$. The remaining seven samples are in poor agreement with the independent age control; the measured to expected dose ratios vary between $1.24 \pm 0.09$ and $7.9 \pm 1.0$. The MAM$\tau$ overestimation tends to increase for decreasing expected dose. If we apply the IEU approach to the multi-grain dose distributions, six of the eight samples have measured to expected dose ratios consistent with unity within two standard errors. Samples T-37 and T-30 significantly overestimate the expected dose; the ratios are $3.9 \pm 0.5$ and $1.5 \pm 0.1$ respectively.
For the extreme case of T-37 where the single grain analysis showed that ~80% of the grains are poorly bleached, the presence of a large number of grains with significant residual doses prevents accurate analysis by any of the statistical analyses used here.

Following the same procedure applied to single grains (see section 5.1.2), we tested the dependence of $D_e$ on additional uncertainty using samples T-39 and T-31. Figure 7b and 7d shows that any additional uncertainty greater than ~7% results in a significant overestimate of the expected dose for both MAM$_{tr}$ and IEU.

6. Discussion

The aim of this study is to explore the most suitable methods for burial dose estimation for flash-flood deposits, and such an investigation requires a set of known age deposits. We consider our age control - historical records and radiocarbon for the oldest sample T-17 (~1000 years) - sufficiently reliable to allow us to assign calendar ages to individual flood events (Benito et al., 2010) and so derive expected equivalent doses.

Single grain dose distributions have been measured from each of our 8 samples, and these distributions analysed using various approaches. Unsurprisingly, those approaches which include all dose points in the estimation (simple average and CAM$_{tr}$) significantly overestimate the expected age, but only by about 1000 years, corresponding to a mean residual dose at deposition of <1 Gy. Thus, despite the fact that our sediments were deposited by short-lived flash floods, the residual doses are comparable with other reports of residuals in modern river sediments, including non-flood deposits. We deduce that most of the light exposure of our sediment samples must have taken place before the final transport by the flash flood which finally deposited the sediment. It is also important to note that although for these
very young samples a dose overestimation of ~1 Gy is unacceptable, such an overestimation would be trivial for older samples (e.g. >10 ka).

The observation that the estimates from robust statistics (which only rejected the upper 15% of grains, on average) are close to the expected ages for three of the samples (T-36, T-23 and T-17) suggests that only a small fraction of the grains were actually incompletely bleached. This is also consistent with the results from the IEU minimum dose approach which identifies these three samples to be the ones with large well-bleached populations, that is 70% (n=78), 77% (n=62) and 75% (n=59) of the grains, respectively. This suggests that if samples containing a moderate percentage of incompletely bleached grains (perhaps <30%) can be identified a priori, a simple non-subjective elimination of high-dose outliers would provide accurate ages; this approach would avoid the need for assumptions concerning OD, and the use of more complex statistical analysis.

The remaining five samples appear to include a higher proportion of incompletely bleached grains in their dose distributions, and minimum age models (IEU and MAM) were necessary to obtain accurate results. All single grain burial dose estimates from the IEU model are in agreement with the doses predicted by the historical record, radiocarbon (and incidentally are completely bracketed by the $^{14}$C modelled ages of Thorndycraft et al., 2012).

The dose distributions from these young samples contain non-positive dose estimates which prevented the direct use of MAM due to its log-normal assumption. However, this problem is easily overcome applying an exponential transformation to the data and then using the original MAM scripts (this analysis using transformed data is termed MAM$_{tr}$). To confirm that ages estimated with MAM$_{tr}$ can be directly compared with any other published results in which estimates were obtained using MAM$_{UL}$, burial doses for all samples have been calculated with both variations of the
original MAM (MAM$_{UL}$ and MAM$_{tr}$) with identical results. In our experience MAM$_{tr}$ is less sensitive to model starting parameters and is more robust in use.

Nevertheless, MAM$_{tr}$ estimates are consistent with the expected ages only for four of the eight samples: the three oldest samples, T-26, T-23 and T-17 (>350 years) and sample T-36 (expected from independent age control to be 119 years and have a corresponding D$_e$ of 0.109 ± 0.006 Gy). This inaccuracy of MAM$_{tr}$ burial dose estimates might be expected to have arisen because the assigned uncertainties (derived from gamma dose recovery experiments) are larger than those actually occurring in nature. However investigation of the sensitivity of MAM$_{tr}$ burial doses to the size of the assigned uncertainty shows that the MAM$_{tr}$ model overestimates the expected doses even for the lowest additional intrinsic OD tested (5%), (see Figure 7); we are confident that the intrinsic OD must be considerably greater than this as the instrument reproducibility alone is ~6% per dose estimate (Thomsen et al., 2005).

It is also interesting to test the performance of the decision tree protocol of Bailey and Arnold (2006) using these data. These authors suggest a decision tree to identify the most appropriate statistical model to apply to a given dose distribution. For our eight samples, this decision tree predicts that the true burial dose is best estimated using the lowest 5% of the dose estimates for all dose distributions except for the oldest sample (T-17) for which the predicted model is MAM-4. The MAM-4 dose estimated for T-17 is 1.03 ± 0.05 Gy, indistinguishable from the MAM$_{tr}$ (3 parameters) result (see Table 2) and consistent with the independent age control. However, for the seven youngest samples (≤450 years old) using the lowest 5% of the dose distribution (after rejection of the negative dose values) results in significant underestimations of the burial doses, of about 80 to 85%.

Despite the expectation that these samples were incompletely bleached, burial doses were also estimated using small multi-grain (~30 grains) aliquots. The
overestimation found when applying simple average and CAMr to these data confirms that if all results are included in the burial dose estimates, then there is a large contribution from incompletely bleached grains giving an average overestimate of <3 Gy (corresponding to ~3000 years). According to the cumulative light sum obtained from single-grain measurements (section 4.2) the total light from a multi-grain aliquot containing ~30 grains must be dominated by the light from ~3.5 grains. Thus at least some of the aliquots can be expected to behave like single grains. In addition IEU results from single grains indicate that 38-77% of the grains from the samples (except T-37) are well bleached. Thus, it is to be anticipated that the signal from a number of multi-grain aliquots will only be derived from well bleached grains. Both minimum age models (MAMr and IEU) were used to estimate the burial dose from multi-grain dose distributions. The MAMr resulted in ages consistent with the known ages only for the two oldest samples, whereas the IEU derived consistent estimates for six of the eight samples. Only the ages of the two youngest samples (T-39 and T-37, 40 and 70 years old, respectively) were overestimated, by approximately 30 and 10 years.

Over the last few years, single grains have been increasingly been used almost routinely for fluvial deposits, despite the measurement and analysis time involved (e.g. Thomas et al., 2005; Arnold et al., 2007; Thomsen et al., 2007). However from the results of this study we conclude that accurate age can be obtained using small (~30 grains) multi-grain aliquots in combination with the IEU approach. In our eight samples this would result in an average offset in dose of $43 \pm 20$ mGy ($n=8$, ~40 years). This suggests that it may not be necessary to resort to single grain analysis of fluvial deposits if deposition ages are much older than a few decades.

Nevertheless, if the IEU model (but not MAMr) is applied to single grain data even this small offset is removed (average offset is $-8 \pm 8$ mGy, $n=8$); this approach gives the most accurate ages overall, including for the youngest samples (<100 years old).
These preferred ages (i.e. obtained using the IEU approach) and the ratio of these ages and those from independent age control are summarized in Table 3.

7. Conclusion

A number of studies have addressed the OSL dating of individual flood layers, most of them from large fluvial catchments (Stokes et al., 2001, Grodek et al., submitted). In small basins (<500 km²) one might expect a large effect from incomplete bleaching in sediments transported by flash-floods over short distances. At the Estrecho site on Guadalentín River slack-water deposits are emplaced by such short-lived flash floods. We have applied a variety of statistical tools in the analysis of single-grain dose distributions from these samples. Unsurprisingly, those models which use all the doses in a distribution (unweighted average and CAM) overestimate the known age. Nevertheless the implied residual doses are small, and are comparable to, or smaller than, reported residuals in a variety of modern river sediments, including non-flood deposits. We conclude that our samples must have been bleached mainly before the final transport by the flash-flood which deposited the sediment. Although small, the residual doses remaining at deposition are significant compared to the doses absorbed since deposition by these known-age young samples, but even an approach which simply rejects outliers in the dose distribution (so-called ‘robust statistics’) is able to provide accurate age estimates in several samples. Of the two models (IEU and MAMₙ) which set out to identify the well-bleached part of the dose distribution, both perform well on older samples, but the IEU is clearly the more accurate for samples containing doses <0.4 Gy (<450 years). Interestingly, acceptably accurate results can be obtained for all but the youngest two samples using the IEU on dose estimates from small multi-grain aliquots, suggesting that the more time-consuming and labour intensive single-grain analyses are not always necessary when analysing incompletely bleached sediments.
Our data show that quartz OSL dating can be used to derive accurate age estimates covering the last few decades to hundreds of years using sediment deposited during flash-floods in small catchments. This opens up the exciting possibility of generating time series of flood discharge at sites without any monitoring or other historical record.

8. Acknowledgments

This study was funded by the Spanish Ministry of Economy and Competitiveness through the research projects FLOOD-MED (ref. CGL2008-06474-C02-01), and CLARIES (ref. CGL2011-29176), and by the CSIC PIE Intramural Project (ref. 200430E595).

9. References


Figure 1. Site location, section and age range association. Position of the eight samples taken for OSL measurements are shown on the section (black dots). Flood year assigned from documentary records and the five radiocarbon dates available are also shown.
Figure 2a. Normalized natural OSL decay curve from a multi-grain aliquot of sample T-31 (solid downwards triangles) and a multi-grain aliquot of calibration quartz (open upwards triangles). The summation intervals used (EBG) are indicated by the dotted lines. The inset shows the sensitivity corrected dose response curve fitted linearly. The recycling point is shown as an open square. The natural (open circle) is shown on the y-axis.
Figure 2b. Normalized natural OSL decay curve from a single grain of sample T-26 (solid downwards triangles) and a single grain of calibration quartz (open upwards triangles). The summation intervals used are indicated by the dotted lines. The inset shows the dose response curve fitted with a saturating exponential function.
Figure 3. Cumulative light sum against the proportion of grains for the natural (dashed line) and the 3 Gy regeneration dose (solid line) signal obtained from the 3600 grains measured from sample T-26. The 1:1 represents the “ideal” case where all grains contribute equally to the light sum.
Figure 4. (a) Preheat plateau results from multi-grain aliquots of samples T-23 (squares) and T-39 (open triangles). The cutheat temperature was 20°C less than the applied preheat temperature. Points and error bars correspond to the average and standard error of 6 individual dose estimates. The overall average and standard error are 1.30 ± 0.12 Gy and 1.19 ± 0.08 Gy for T-23 (solid line) and T-39 (dashed line), respectively. All dose estimates are consistent with the overall average within 2 standard errors showing no significant offsets at higher temperatures. (b) Thermal transfer results from laboratory bleached aliquots of samples T-17 (triangles), T-26 (squares) and T-39 (open circles). Individual points correspond to the average of 6 independent dose estimates for each temperature. Thermal transfer only appear to become significant for temperatures above 240°C. (c) CAM_tr results from thermal transfer laboratory blue light bleached single grains from sample T-30 (upwards solid triangles, ~30 grains per dose estimate). The dose derived from solar simulator (Hönle Sol 2) bleached single grains of sample T-39 using a preheat temperature of 200°C is also shown (downwards open triangle). The inset shows the dose distribution obtained for the bleached portion of sample T-39 using a preheat temperature of 200°C.
Figure 5. Single grain dose distributions from the eight samples measured using the SAR protocol. We show the signal of the natural test dose as a function of measured dose. The insets show the same data for doses less than 1.3 Gy except for T-17 where doses less than 1.9 Gy are shown. The expected doses are indicated by the vertical red line/bar. These expected doses are derived from the ages provided by the independent age control multiplied by the measured quartz dose rate given in Table 1.
Figure 6. Estimated doses using (a) single grains and (b) multi-grain aliquots versus expected doses. Doses have been estimated using the simple average (solid squares), the robust average (open squares), CAM$_{tr}$ (solid orange circles), MAM$_{tr}$ (upward gray triangles) and IEU (downward blue triangles). Insets show a close up of the burial dose estimates for the seven youngest samples using MAM$_{tr}$ and IEU. The 1:1 line (dashed) is also shown.
Figure 7. Ratio of estimated minimum burial dose and the expected dose as a function of additional uncertainty assigned to individual dose estimates for sample T-39 (a and b) and T-31 (c and d) using MAM$_{tr}$ (squares) and IEU (triangles). (a) and (c) are derived from measurements using single grains; (b) and (d) are derived from multi-grain measurements. The open symbols indicate the burial dose estimates obtained using the additional uncertainty value derived from the gamma dose recovery experiments.
Figure 8. Multi-grain aliquot dose distributions from the eight samples measured using the SAR protocol. The signal of the natural test dose is plotted as a function of the measured dose. The insets show the same data from doses less than 1.8 Gy except for T-17 where doses less than 3.2 Gy are displayed. The expected doses are indicated by the vertical red line/bar.
Table 1. Summary of sample depth, estimated water content and radionuclide concentrations used to calculate the total quartz dose rate. The water content was measured immediately after sampling and is considered representative for the burial time. An absolute error of 4% is assumed. Uncertainties represent one standard error. An internal quartz dose rate of 0.06 ± 0.03 Gy/ka has been assumed. No uncertainties are reported for the expected ages as they are based on detailed historical records, e.g. they are known to a precision of days. The exception to this is sample T-17 where the expected age is based on radiocarbon dating (see text for details).
<table>
<thead>
<tr>
<th>Sg. (Single-grains)</th>
<th>T-39</th>
<th>T-37</th>
<th>T-36</th>
<th>T-31</th>
<th>T-30</th>
<th>T-26</th>
<th>T-23</th>
<th>T-17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected (Gy)</td>
<td>0.033 ± 0.002</td>
<td>0.062 ± 0.003</td>
<td>0.109 ± 0.006</td>
<td>0.15 ± 0.01</td>
<td>0.18 ± 0.01</td>
<td>0.32 ± 0.02</td>
<td>0.37 ± 0.02</td>
<td>1.10 ± 0.12</td>
</tr>
<tr>
<td>Av.</td>
<td>3.8 ± 0.8</td>
<td>3.7 ± 0.6</td>
<td>87</td>
<td>1.4 ± 0.3</td>
<td>80</td>
<td>1.3 ± 0.4</td>
<td>94</td>
<td>5.0 ± 1.3</td>
</tr>
<tr>
<td>Norm. Av.</td>
<td>114 ± 31</td>
<td>60.0 ± 13.4</td>
<td>12.4 ± 3.0</td>
<td>9.0 ± 2.9</td>
<td>28.1 ± 8.6</td>
<td>11.2 ± 2.3</td>
<td>10.3 ± 3.6</td>
<td>4.5 ± 1.0</td>
</tr>
<tr>
<td>Robust av.</td>
<td>1.7 ± 0.2</td>
<td>67</td>
<td>1.9 ± 0.2</td>
<td>76</td>
<td>0.63 ± 0.07</td>
<td>69</td>
<td>0.52 ± 0.05</td>
<td>77</td>
</tr>
<tr>
<td>Norm. Robust av.</td>
<td>51 ± 10</td>
<td>31.4 ± 4.0</td>
<td>5.8 ± 1.0</td>
<td>3.6 ± 0.6</td>
<td>6.7 ± 1.3</td>
<td>5.5 ± 0.9</td>
<td>2.0 ± 0.3</td>
<td>3.2 ± 0.6</td>
</tr>
<tr>
<td>CAM&lt;sub&gt;r&lt;/sub&gt;</td>
<td>3.5 ± 0.7</td>
<td>90</td>
<td>3.5 ± 0.5</td>
<td>87</td>
<td>1.2 ± 0.2</td>
<td>80</td>
<td>1.3 ± 0.3</td>
<td>94</td>
</tr>
<tr>
<td>Norm. CAM&lt;sub&gt;r&lt;/sub&gt;</td>
<td>107 ± 28</td>
<td>56.1 ± 11.2</td>
<td>10.7 ± 2.4</td>
<td>8.6 ± 2.7</td>
<td>28.7 ± 8.4</td>
<td>10.4 ± 2.0</td>
<td>8.1 ± 2.4</td>
<td>3.3 ± 0.7</td>
</tr>
<tr>
<td>OD (%)</td>
<td>176 ± 19</td>
<td>139 ± 17</td>
<td>145 ± 20</td>
<td>245 ± 29</td>
<td>207 ± 25</td>
<td>126 ± 15</td>
<td>234 ± 29</td>
<td>67 ± 11</td>
</tr>
<tr>
<td>MAM&lt;sub&gt;r&lt;/sub&gt;</td>
<td>0.26 ± 0.03</td>
<td>39</td>
<td>0.35 ± 0.06</td>
<td>14</td>
<td>0.19 ± 0.02</td>
<td>38</td>
<td>0.27 ± 0.02</td>
<td>55</td>
</tr>
<tr>
<td>Norm. MAM&lt;sub&gt;r&lt;/sub&gt;</td>
<td>7.8 ± 1.3</td>
<td>5.7 ± 1.2</td>
<td>1.8 ± 0.3</td>
<td>1.9 ± 0.2</td>
<td>2.1 ± 0.3</td>
<td>1.5 ± 0.2</td>
<td>1.2 ± 0.1</td>
<td>0.9 ± 0.3</td>
</tr>
<tr>
<td>IEU</td>
<td>0.06 ± 0.04</td>
<td>11</td>
<td>0.24 ± 0.03</td>
<td>5</td>
<td>0.11 ± 0.01</td>
<td>17</td>
<td>0.16 ± 0.01</td>
<td>25</td>
</tr>
<tr>
<td>Norm. IEU</td>
<td>1.9 ± 1.4</td>
<td>3.9 ± 0.7</td>
<td>1.0 ± 0.2</td>
<td>1.1 ± 0.1</td>
<td>1.5 ± 0.2</td>
<td>1.1 ± 0.1</td>
<td>1.0 ± 0.1</td>
<td>1.0 ± 0.2</td>
</tr>
</tbody>
</table>

Table 2. Summary of burial doses for both multi-grain small aliquots (SA) and single-grains (SG). “Expected” is the dose expected from independent age control. The uncertainties assigned to the expected dose is derived from the uncertainties assigned to the respective dose rates. The uncertainty assigned to sample T-17 also contains the uncertainty assigned to the radiocarbon date. “Av.” is the simple unweighted average and “Robust av.” is that derived from robust statistics. CAM<sub>r</sub> and MAM<sub>r</sub> are the transformed version of the CAM and MAM approaches respectively. OD is the over-dispersion derived from CAM<sub>r</sub>. IEU is the internal/external consistency criterion. The ratio between each result and the dose expected from the independent age control are summarized as “Norm.”. The number of aliquots (n) used in each case is also given.
<table>
<thead>
<tr>
<th>Sample</th>
<th>Depth (cm)</th>
<th>Preferred ages</th>
<th>Expected ages</th>
<th>Estimated/Expected age ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-39</td>
<td>59</td>
<td>0.07 ± 0.02</td>
<td>0.040</td>
<td>1.8 ± 0.4</td>
</tr>
<tr>
<td>T-37</td>
<td>95</td>
<td>0.08 ± 0.02</td>
<td>0.069</td>
<td>1.2 ± 0.3</td>
</tr>
<tr>
<td>T-36</td>
<td>122</td>
<td>0.11 ± 0.01</td>
<td>0.119</td>
<td>0.90 ± 0.11</td>
</tr>
<tr>
<td>T-31</td>
<td>216</td>
<td>0.16 ± 0.02</td>
<td>0.172</td>
<td>0.92 ± 0.10</td>
</tr>
<tr>
<td>T-30</td>
<td>259</td>
<td>0.20 ± 0.02</td>
<td>0.179</td>
<td>1.12 ± 0.09</td>
</tr>
<tr>
<td>T-26</td>
<td>351</td>
<td>0.31 ± 0.03</td>
<td>0.357</td>
<td>0.85 ± 0.08</td>
</tr>
<tr>
<td>T-23</td>
<td>367</td>
<td>0.40 ± 0.03</td>
<td>0.442</td>
<td>0.90 ± 0.07</td>
</tr>
<tr>
<td>T-17</td>
<td>478</td>
<td>0.96 ± 0.06</td>
<td>0.96 ± 0.06</td>
<td>0.98 ± 0.06</td>
</tr>
</tbody>
</table>

Table 3. Summary of the preferred ages and the ratio of these ages and the ages expected from independent age control. The preferred ages are based on the burial doses estimated using the IEU approach.