1-1-2012

Realizing the potential of fluvial archives using robust OSL chronologies

Alastair C. Cunningham
Delft University of Technology, acunning@uow.edu.au

Jakob Wallinga
Delft University of Technology

Realizing the potential of fluvial archives using robust OSL chronologies

Abstract

Optically Stimulated Luminescence (OSL) dating has enormous potential for interpreting fluvial sediments, because the mineral grains used for OSL dating are abundant in fluvial deposits. However, the limited light exposure of mineral grains during fluvial transport and deposition often leads to scatter and inaccuracy in OSL dating results. Here we present a statistical protocol which aims to overcome these difficulties. Rather than estimating a single burial age for a sample, we present ages as likelihood functions created by bootstrap re-sampling of the equivalent-dose data. The bootstrap likelihoods incorporate uncertainty from age-model parameters and plausible variation in the input data. This approach has the considerable advantage that it permits Bayesian methods to be used to interpret sequences containing multiple samples, including partially bleached OSL data. We apply the statistical protocol to both single-grain and small-aliquot OSL data from samples of recent fluvial sediment. The combination of bootstrap likelihoods and Bayesian processing may greatly improve OSL chronologies for fluvial sediment, and allow OSL ages from partially bleached samples to be combined with other age information.

Keywords

CAS

Publication Details

Cunningham, A. C. & Wallinga, J. (2012). Realizing the potential of fluvial archives using robust OSL chronologies. Quaternary Geochronology, 12 (October), 98-106.

Realizing the potential of fluvial archives using robust OSL chronologies

Alastair C. Cunningham*[1], Jakob Wallinga[1]

[1]*Netherlands Centre for Luminescence dating, Delft University of Technology, Faculty of Applied Sciences, Mekelweg 15, NL-2629 JB, Delft, The Netherlands*

*Corresponding Author:

Email: blinkysimpson@hotmail.com

Tel: +31 1527 86619

Fax: +31 1527 86422

Abstract

Optically Stimulated Luminescence (OSL) dating has enormous potential for interpreting fluvial sediments, because the mineral grains used for OSL dating are abundant in fluvial deposits. However, the limited light exposure of mineral grains during fluvial transport and deposition often leads to scatter and inaccuracy in OSL dating results. Here we present a statistical protocol which aims to overcome these difficulties. Rather than estimating a single burial age for a sample, we present ages as likelihood functions created by bootstrap re-sampling of the equivalent-dose data. The bootstrap likelihoods incorporate uncertainty from age-model parameters and plausible variation in the input data. This approach has the considerable advantage that it permits Bayesian methods to be used to interpret sequences containing multiple samples, including partially bleached OSL data. We apply the statistical protocol to both single-grain and small-aliquot OSL data from samples of recent fluvial sediment. The combination of bootstrap likelihoods and Bayesian processing may greatly improve OSL chronologies for fluvial sediment, and allow OSL ages from partially bleached samples to be combined with other age information.

Keywords

OSL, luminescence, bleaching, fluvial, Bayesian, OxCal, chronology, bootstrap, likelihood

1. Introduction

Sedimentary deposits of river-transported material provide an important record of environmental history. Fluvial sediments are widely studied to understand modern fluvial sedimentation rates (e.g. Owens et al., 1999; Hobo et al., 2010), determine fluvial response to climatic, tectonic and sea-level forcing (e.g. Busschers et al., 2008), and to reconstruct flood risks (e.g. Benito et al., 2008). However, the use of fluvial archives is severely hindered by the lack of consistent dating. Accurate and precise dating is clearly essential for correlating fluvial sedimentation with external forcing. Fluvial sediments are non-continuous and lack the annual layering necessary for high-precision methods; dating control must therefore be obtained through radiometric methods. Radiocarbon dating offers the most precision, but is of limited use for direct dating of fluvial activity due to the frequent absence of organic carbon, and because the carbon is often re-worked from older deposits. In contrast, Optically Stimulated Luminescence (OSL) dating is nearly always possible, because the raw material for OSL dating - sand-sized mineral grains - is abundant in fluvial sediment. OSL dating also has the advantage of a wide age-range of applicability (~10 a to >100 ka). With these advantages, OSL dating could provide continuity in a multi-dating-method chronology, and become the standard method for dating fluvial sediment (Wallinga, 2002; Rittenour, 2008).

OSL dating requires determination of the radiation dose absorbed by the mineral grains since burial (the burial dose), and the radiation dose rate. It is the determination of the burial dose that presents difficulties in dating fluvial sediment. The problem lies with the most fundamental requirement for obtaining an age with OSL techniques – that the mineral grains were exposed to enough sunlight during the last episode of transport and deposition for the OSL signal to be reset. A few tens of seconds of bright sunlight is enough for resetting, but the equivalent light exposure is

not always received by grains transported within the water column. The effect is usually known as ‘partial bleaching’ (or ‘heterogeneous bleaching’), with the consequences for age determination dependent on the severity of the effect. Partial bleaching tends to be most problematic where deposition is more recent (e.g. within the last 2000 a, Jain et al., 2004).

It is notable that where OSL has proven successful in interpreting fluvial systems (e.g. Rittenour et al., 2005; Rodnight et al., 2005; Busschers et al., 2007), the degree of partial bleaching in the data is minimal. The appearance of partial bleaching in a dataset necessitates some statistical processing, although the selection and application of ‘age models’ is a frequent source of discussion (Bailey and Arnold, 2006; Rodnight et al., 2006; Arnold and Roberts, 2009; Thrasher et al., 2009). Difficulties arise due to the sensitivity of the burial dose to the lowest D_e value, which may or may not be an outlier, and in assessing the amount of spread in the data that can be assigned to the burial-dose population. As there is no commonly agreed procedure for coping with these issues, there is a degree of inconsistency in age-model application. More devastatingly, the error terms assigned to the burial ages reflect (at best) the uncertainty in fitting the model to the data, and take no account of uncertainty in the decision process itself. As a consequence, OSL ages for fluvial sediments often appear scattered or inaccurate, with error terms that are less than meaningful.

The aim of this paper is to provide a robust protocol for the analysis of OSL data from fluvial (or glaciofluvial) sediment. We use both single-grain and small-aliquot data from a fluvial sequence, allowing us to test the validity of using multi-grain aliquots for partially bleached samples. We show that by embedding partially bleached OSL data in a Bayesian framework, the coherence of an OSL chronology can be increased. The use of Bayesian methods requires the construction of a likelihood function for the OSL age, for which we develop a new method based on bootstrap re-sampling of the D_e distribution. The method is able to incorporate uncertainties in the D_e distribution and age-model parameters, and through the combination with Bayesian statistics leads to an objective means of identifying outliers.

2. Methods

2.1 Sample details

We use a sequence of seven OSL samples taken from a single core through embanked floodplain sediments of the River Waal, The Netherlands. The sediments were deposited over the last 1000 years; the OSL data show far more overdispersion than would be expected from well-bleached samples (Table 1). Some bioturbation of the upper part of the sequence can be expected, as some smoothing of heavy metal profiles has been observed (Hobo et al., 2010). In the lower part of the core, a rapid rate of deposition is likely to have precluded this effect. For most of the relevant time period, there is no alternative dating method available for these sediments. OSL measurements were performed firstly on multi-grain aliquots of 100-200 grains each, with details described in Wallinga et al. (2010); additional site information and alternative dating methods are presented by Hobo et al. (2010). For the current paper we include three additional samples of the underlying channel deposits taken from the same core; all OSL decay curves were re-analysed using the 'early background' subtraction described in Cunningham and Wallinga (2010). Integration intervals were 0-0.4 s for the initial signal, 0.4-1.4 s for the background, under ~40 mW cm^{-2} blue LED stimulation.

New single-grain measurements were performed on all 7 samples, using a Risø TL/OSL-DA-15 reader with single-grain attachment (Bøtter-Jensen et al., 2000). Single-grains were optically stimulated using an Nd:YVO_4 diode-pumped laser (λ= 532 nm). The detection filter was a 2.5 mm Hoya U340, following Ballarini et al. (2005). The natural and test-dose OSL was measured for all grains; grains with a relative standard error on the first test-dose OSL of less than 6.5% were selected for the complete measurement protocol, with other grains ignored in the analysis. Signal analysis for single grains also used the early background subtraction (0-0.17 s for the initial signal, 0.17-0.58 s for the background). The measurement protocol for single

grains was otherwise identical to the multi-grain protocol. Grains were accepted if their recycling ratios were between 0.9 and 1.1, and if recuperation was less than 10% of the regenerative dose.

Table 1 about here

2.2 Bayesian chronological framework

Bayesian methods have long been recognised as a powerful aid in the analysis of age information (Buck et al., 1991, Bronk Ramsey, 1995). A Bayesian chronological framework has two particular uses: it provides a formal method of combining multiple age estimates into a meaningful chronology (including an objective means of identifying outliers), and it utilises stratigraphic relationships between the samples to increase dating precision. Bayesian methods have gained widespread use with radiocarbon-based chronologies (e.g. Blockley et al., 2007, Jacobi and Higham, 2009, Bronk Ramsey et al., 2010), where the analysis helps discriminate between multiple peaks in calibrated age probability distributions.

The power in Bayesian techniques comes through the incorporation of 'prior' information, i.e. information known before measurement of any sample. For sedimentary sections, this comes from the stratigraphic relationship between the sample locations, which may simply constrain the order in which the samples were deposited, or may contain more detailed assumptions about the depositional process (Bronk Ramsey, 2008). The chronological model is developed through the combination of the prior model with the age information obtained from measurements (the 'likelihood'), input in the form of a probability density function (PDF).

Given the ability of Bayesian analysis to identify outliers and increase precision, it is clearly of interest in processing OSL ages derived from heterogeneously bleached samples. There are a number of freely available chronological tools that make use of Bayesian statistics (see Parnell et al., 2011). One such program is OxCal (Bronk Ramsey, 1995), which is widely used for analysis of

radiocarbon dated sequences, and can also be used to include age information from other methods (e.g. OSL ages from well-bleached samples; Rhodes et al., 2003). The wholesale inclusion in OxCal of a sequence of fluvial OSL samples has not yet been attempted, and this could be a reflection of inaccuracy or spurious precision in ages assigned to fluvial samples.

OxCal requires age information in the form of a PDF. For a well-bleached OSL sample with the age defined with a 1σ error term, this is easily achieved using the internal functions of OxCal (see Rhodes et al. (2003) for details). For partially bleached samples, the creation of a PDF is not so straightforward: the OSL age may be dependent on the age model used and the assumptions that go with that model, and the use of a normally distributed error term may not be valid. What is required, therefore, is a means of estimating a likelihood function for the age of a sample, incorporating the different sources of error. In the sections that follow, we show how bootstrap methods can be used to create an analogue of the likelihood function.

2.3 Bootstrap likelihoods

Outline of procedure

Measurements of equivalent dose (D_e) can be made on single grains or on multi-grain aliquots. In either case, we can define a dataset of $x = (x_1,x_2,...,x_n)$ of n D_e estimates. Each $x_i = (y_i,s_i)$, that is, each x_i consists of an estimate of D_e (y_i), and an estimate of the standard error of that measurement (s_i). We wish to estimate θ, the mean radiation dose received by the grains since they were last buried (the 'burial dose'); our estimate of θ is denoted . The age of the sample is then estimated by / , where is the mean dose rate to the grains. For partially bleached samples, a commonly used method of calculating is using the 3-component minimum-age model (MAM3) of Galbraith et al. (1999). Under this model, the parameter $\gamma = \log(\theta)$ is estimated using a maximum likelihood approach. The $\log(x_i)$s are assumed to belong to a population equal to γ, or to a second population greater than γ represented by a half a normal distribution. With this model, it is assumed that the dispersion in the population of well-bleached grains

is entirely accounted for by the associated error terms. The evidence from dating of well-bleached (e.g. aeolian) sediment consistently shows that this assumption is not reasonable, and so Galbraith et al. (2005) introduced the term σ_b to the age models. σ_b can be included in the MAM3 by increasing the s_is (Galbraith and Roberts, in press), effectively allowing the well-bleached population to be defined by a log-normal distribution with mean and relative standard deviation σ_b.

In this paper we use an altered, 'unlogged' version of the MAM3 described by Arnold et al. (2009), henceforth the $MAM3_{ul}$. This unlogged version is more suitable for very young sediments as it can deal with estimates that are equal to zero within their uncertainty limits. Rather than calculating a single estimate of θ, we design a protocol for creating a probability density function to represent the likelihood as a function of θ. The protocol can be described as a bootstrap partial likelihood, and is summarised below. Each step of the protocol is expanded upon in the following subsections.

Bootstrap likelihood protocol:

1. Create a bootstrap sample from the original data
2. Stochastically generate σ_b
3. Calculate the bootstrap replicate with likelihood estimated using a nested bootstrap or bootstrap recycling.
4. Incorporate unshared systematic error.
5. [after repeating steps 1-4 many times] Apply polynomial smoothing to the pairs of [θ, $L(\theta)$].

Bootstrap resampling

The bootstrap was introduced by Efron (1979) as a non-parametric means of estimating the standard error of the parameter of interest. A full account can be found

in Efron & Tibshirani (1993). With this method, a bootstrap sample $x^* = (x^{*1}, x^{*2}, \ldots x^{*n})$ is drawn by random sampling with replacement from the original dataset x, of length n. This process is performed repeatedly, with each bootstrap sample used to create a bootstrap replicate The function $s(\cdot)$ is the same as that applied to the original data x, in our case the $MAM3_{ul}$.

The σ_b parameter

When applied to the minimum-age models, σ_b represents the overdispersion in the data that would be expected should the sample of interest be well-bleached. It is a fixed parameter of the minimum-age models, and must be estimated before a model is run. An overestimate of σ_b will lead to an overestimate of the burial dose (and hence the age), an underestimate in σ_b will lead to an underestimate the burial dose and age.

It is far from certain what the value of σ_b should be, and it is likely to be sample dependent. The influences on σ_b can be categorised as follows:

- Errors arising during measurement – different grains may react differently to optical and thermal stimulation, causing them to yield different D_e; see Thomsen et al. (2005; 2007).

- Grain-to-grain variation in the dose rate received by grains in nature. This could arise through the localised concentrations of beta sources in sediment (e.g. feldspars or zircons, Mayya et al., 2006), or through the presence of macro bodies of non-radioactive material (Nathan et al., 2003, Cunningham et al., 2011a).

- Calculation of measurement errors. Because σ_b estimates the spread in the data beyond that caused by the s_is, it is dependent on the way the s_is are calculated. We could therefore expect σ_b to be dependent on the laboratory which produced the data, as methods of calculating s_i vary between laboratories.

Ideally, the expectation of σ_b would be calculated on a sample-by-sample basis. In the absence of such information, a value of 0.20 (i.e., 20% overdispersion) could be a respectable approximation at the single grain level. This value is the mean overdispersion from a large number of single-grain studies on well-bleached samples Arnold and Roberts (2009). However, Arnold and Roberts (2009) also showed a significant amount of variation exists between different samples. In the bootstrap likelihood protocol presented here, uncertainty in σ_b can be incorporated in the likelihood profile by including stochastic variation in σ_b. For each bootstrap sample x^*, a value of σ_b is drawn randomly from a normal distribution; we use a normal distribution with mean of 0.20 and standard deviation of 0.04 for the single-grain data. For multi-grain data, σ_b must be smaller than for single-grain data from the same sample. When there is more than one grain in an aliquot, grain-to-grain variation in D_e will tend to get averaged, reducing the overdispersion for a well-bleached sample. This process has been modelled by Cunningham et al. (2011b), who found that the extent of the averaging effect is dependent on the number of grains in the aliquot and the single-grain sensitivity distribution of the sample. Following the protocol of Cunningham et al. (2011b), σ_b for the multi-grain data in this study is estimated to be 0.11 ± 0.04.

Likelihood estimates

Having obtained a bootstrap replicate by running the $MAM3_{ul}$ with a bootstrap sample and stochastically generated σ_b, it is necessary to associate a likelihood with that value. The bootstrap partial likelihood approach estimates this with a nested bootstrap calculation (Davison et al., 1992; Efron and Tibshirani, 1993). From each of M bootstrap samples x^{*i}, (i=1:M), we generate N second-level bootstrap samples by sampling with replacement from x^{*i}. The likelihood at is estimated using a kernel density estimate of the second-level bootstraps:

$$L(\theta) = \frac{1}{Nh}\sum_{m=1}^{N} k\left(\frac{t-\hat{\theta}_m^{**}}{h}\right)$$

where $k(\cdot)$ is the kernel density estimate with bandwidth h, and where is a second-level bootstrap replicate. We use the standard normal kernel, with bandwidth h determined by the standard error on when using the original data x. The bandwith is therefore roughly proportional to the age of the sample, and wider when is uncertain.

A consequence of nested bootstrap calculations is a large computational burden. The total number of θ evaluations is $M(N+1)$. For our purposes, reasonable values for M and N are about 2000 and 100, respectively, leading to ~200,000 calls to the $MAM3_{ul}$. Because each evaluation of the $MAM3_{ul}$ is relatively expensive, the total computational burden is prohibitive. The bootstrap recycling procedure (Newton and Geyer, 1994) was developed to solve this problem, and is outlined succinctly by Davison et al. (1995). Rather than sampling the second-level bootstraps from each first-level bootstrap sample, they are drawn from one probability vector p^0 of the original sample x. Weights are used to achieve the same effect as sampling from the first-level bootstrap probability vector p^*. The likelihood equation under bootstrap recycling becomes

$$L(\theta) = \frac{1}{Nh}\sum_{m=1}^{N} k\left(\frac{t-\hat{\theta}_m^{**}}{h}\right)\prod_{j=1}^{n}\left(\frac{P_j^*(\theta)}{P_j^0}\right)^{f_{jm}^{**}}$$

where is the frequency of data value x_i in the m^{th} sample drawn from p^0.

Using bootstrap recycling, the total number of θ evaluations is reduced to $M+N$, although the value of N should be much greater than with the nested bootstrap. The end result is a series of M bootstrap replicates of , each paired with a likelihood estimate $L(\theta)$.

Unshared systematic error

There are sources of uncertainty in OSL dating that are systematic between aliquots, but random between samples. For example, a random error in the dose-rate measurement, or water content correction, will affect all aliquots within a single sample in the same way. Following Rhodes et al. (2003), we refer to this sort of error as unshared systematic (USS) error. However, while Rhodes et al. (2003) used the agreement within the chronological model to determine the USS, we prefer to estimate the USS independently. The USS is incorporated by randomising each bootstrap replicate , using a normal distribution with mean of and standard deviation 0.035 (i.e. 3.5% USS).

Polynomial smoothing

The final step is to fit a smooth likelihood curve through the pairs [, L(θ)]. We use a polynomial function, which provides a reasonable fit (Fig. 1), although more advanced methods could also be used. The fitting is performed on the logged data to homogenize variability.To make full use of the likelihoods, the x-axis needs to be converted from dose to age. This can be done at any stage using the dose rate. This conversion implies that systematic uncertainty in the dose rates should be considered. However, since this would apply to all samples in the same way, it should be included after the chronological model has been constructed (but before comparison with independent ages).

Figure 1 about here

The bootstrap likelihood protocol described above does not produce a true likelihood: a function that is proportional to the probability of a fixed event in sample space (Efron & Tibshirani, 1993). The bootstrap likelihood is an analogue of a partial likelihood, with which it is possible to combine prior information using Bayes' Theorem (Davison et al., 1992). This combination is demonstrated in section 3.

3. Results

Figure 2 about here

We have applied the bootstrap likelihood protocol to the sequence of young fluvial samples. The resulting age distributions are informative, and are shown in stratigraphical order in Fig. 2, along with the profile likelihood of the $MAM3_{ul}$ age. For all samples, the bootstrap likelihoods are broader than the $MAM3_{ul}$ profile likelihood, due to the inclusion of additional sources of uncertainty. The $MAM3_{ul}$ is somewhat sensitive to the lowest precise D_e, resulting in non-normal or multi-modal bootstrap likelihoods. This sensitivity is picked up by the bootstrap likelihoods because some of the bootstrap samples do not contain the lowest D_e value.

Using OxCal v4.1 (Bronk Ramsey, 2008, 2009), the bootstrap likelihoods can be used to create a coherent chronology for the fluvial sediment. The likelihood functions were saved as text files in the OxCal directory, with the units as years AD, and the file suffix *prior*. OxCal provides a number of depositional models and constraints to help define the chronology. We used the *P_Sequence* mode of deposition, as it is most consistent with non-continuous floodplain deposition; we assumed an average of 10 depositional events per metre. We also included a *Tau_Boundary* at the top of the sequence, which formulates a prior model for an exponentially decreasing floodplain sedimentation rate over time. The validity of the sedimentation-rate model is discussed later, the precise command list was:

```
Plot()
 {
  P_Sequence("Site1107",10)
  {
   Boundary("b_old");
   Prior(Sample7){ z=9.42; };
   Prior(Sample6){ z=3.55; };
```

```
   Prior(Sample5){ z=1.64; };
   Prior(Sample4){ z=1.26; };
   Prior(Sample3){ z=0.77; };
   Prior(Sample2){ z=0.31; };
   Prior(Sample1){ z=0.13; };
   Tau_Boundary("b_young");
  };
 };
```

where e.g. 'Sample1' corresponds to a file named 'Sample1.prior' containing the bootstrap likelihood. We are aware that OxCal terminology used here may be confusing: the 'prior' files contain the measurement data and not the prior information on e.g. depth and order of the samples. After running the model, OxCal produces a new series of PDFs, referred to as *Posteriors*. These have been plotted according to depth in Fig. 3 (for single-grain data) and Fig. 4 (for multi-grain data), along with the likelihoods. OxCal also determines an 'agreement index' for each sample (Table 1), and for the overall model. The agreement index gives an objective score of the overlap between the modelled posteriors and the likelihoods. It is suggested that a lower threshold of 60% should be applied to the samples, i.e. data should be rejected if the agreement index for the sample is below 60% (Bronk Ramsey, 2008). For both the single-grain and multi-grain datasets, sample 5 gave an agreement score far below 60%; the age models plotted in Figs 3 and 4 omit this sample.

Figures 3 and 4 about here

4. Discussion

4.1 Advantages of using bootstrap likelihoods

The potential benefit of using Bayesian methods for fluvial sediments is large, but rests on a number of basic assumptions. The first of these is that the likelihood distribution is a good reflection of the uncertainty associated with the OSL measurements. If the likelihood distribution is too narrow, then the lack of coherence between the samples will make it difficult to fit a depositional model; too broad and the model will tend towards more uniform rate of deposition.

The bootstrap routine presented here provides a robust estimation of the minimum-age uncertainty. By testing the sensitivity of the minimum age to (plausible) variation in the input data, the width of the probability distribution is made dependent on the quality of the original data. In a given sequence of fluvial samples, it is probable that some samples will appear better bleached than others. With the Bayesian procedure described above, it should be possible to 'anchor' the chronology on these better-bleached samples.

The application of Bayesian statistics requires careful consideration of the sources of error. In the model discussed so far, systematic errors that are shared between the samples are not included, and must be added to the final (post-OxCal) age estimates. If independent age information is included in the deposition model, then the shared systematic errors should be added before the Bayesian modelling. However, the likely size of shared systematic errors (< 5%) may be insignificant compared to the width of the bootstrap uncertainty distributions.

4.2 Validity of parameters used in the chronological model

OxCal offers a variety of parameters which can be used to specify the prior information about the sedimentation process. The prior information that we have comes from principles of the sedimentation process on embanked floodplains. Sedimentary chronologies are ordinarily based on a *P_Sequence* model, which constrains each model iteration to appear in depth order, while allowing slight

variation in the sedimentation rate between samples. The degree of variation in sedimentation rate is governed by the parameter *k*, which specifies the average number of deposition events per depth unit. We used a low value of k=10, reflecting the sporadic distribution of deposition events (floods) over time. A Tau_Boundary at the top of the stratigraphic model forces a decreasing sedimentation rate, reflecting the reduction in accommodation space as the floodplain builds up.

A different choice of model parameters would lead to different posterior distributions. In particular, a higher *k* would lead to a more uniform model with lower agreement scores, but with the sedimentation rates largely the same. The purpose of using OxCal here is to demonstrate the potential of the bootstrap likelihoods; we have avoided sample rejection to facilitate comparison between single-grain and small-aliquot data.

4.3 Single grain or small aliquots?

There is a great deal of similarity between the inferred ages from single-grain and small-aliquot data, both in the bootstrap likelihoods and the posterior distributions. For samples 1, 4 and 7, the bootstrap likelihoods are similar for both datasets. Sample 5 produces an imprecise, bimodal likelihood for the single-grain data, and is in poor agreement with the rest of the chronology. In the single-grain data, the likelihood for sample 6 is also imprecise, and also has weak agreement with the inferred chronology.

The similarity of the two datasets conflicts with the received opinion that multi-grain aliquots can not be used to date partially bleached sediment. The argument for this is that averaging of the signal from different grains occurs when the OSL is measured on a multi-grain aliquot; a single, poorly bleached grain can therefore corrupt the whole aliquot. What is missing from this argument is an appreciation of the spread in OSL sensitivity between different grains. The OSL sensitivity varies dramatically between grains, and the sensitivity distribution varies dramatically between samples. Differences in sensitivity could reflect different crystal

characteristics, or sensitivity changes brought about through irradiation and bleaching. Quartz grains that have undergone repeated cycles of bleaching and deposition tend to become sensitized (e.g. Pietsch et al., 2008). As a consequence, for some samples a large fraction of quartz grains will yield a measurable OSL signal. For these samples, single-grain dating is efficient, because a significant fraction of the single-grain measurements provide useful data. If this type of sample is partially bleached then single-grain dating is essential. Small aliquots will contain many sensitive grains, leading to a high degree of averaging across the aliquot.

In many locations, sensitivity of the quartz is far less ideal. Samples from any environment can show poor sensitivity (e.g. Fitzsimmons, 2011; Lukas et al., 2007), and highly-skewed sensitivity distributions (Duller, 2008). It is not uncommon for 95% of the combined OSL signal to come from less than 5% of the grains. In our experience of dating quartz from the Netherlands, a single-grain disc of 100 grains typically contains about 1 or 2 sensitive grains. In a multi-grain aliquot of 100 grains, the number of bright grains on the disc can be estimated from the binomial distribution (with n=100 and p=0.015 in this case). For such samples, single-grain dating is *very* inefficient, because the vast majority of single-grain measurements are discarded. Furthermore, single-grain dating is not necessary for partially bleached samples of this type; a small aliquot contains very few sensitive grains, so the averaging effect will be weak.

For the present study, roughly 25700 single grains were initially measured, of which 340 grains (1.3%) were considered sensitive enough to be worth completing the measurements. Only 133 grains (0.5%) passed the acceptance criteria. For multi-grain aliquots, 45% of the measurements yielded D_e values which passed the acceptance criteria. Given the similarity of results, and the greater efficiency of the multi-grain aliquot measurements, we can see little benefit in using currently available single-grain measurement protocols for samples such as these. Nevertheless, the averaging effect will always be present in small-aliquot data. The aliquot size should be restricted as much as feasible, with single grain measurements performed if the sensitivity distribution permits. A discussion on the averaging effect can be found in Duller (2008) and Cunningham et al. (2011b).

4.4 Implications for sampling

The combination of bootstrap uncertainty distributions with Bayesian chronological modelling has the potential to greatly increase the accuracy and precision in dating fluvial deposits. However, for this potential to be realised there are two important requirements of the sampling strategy:

1. *High-resolution sampling*. The use of Bayesian statistics is only beneficial when the uncertainty distributions of different samples overlap. It is therefore essential that sampling resolution is high.

2. *Collection of high-quality stratigraphic information*. The more prior information that can be incorporated into the Bayesian modelling, the greater the precision of the chronological model.

The importance of these points can be seen by considering the chronological model in Fig. 4. In the lower part of the sequence, the posterior distributions are almost identical to the prior distributions, because the poor sampling resolution has lead to prior likelihood distributions that do not overlap.

6. Conclusions

Bootstrap re-sampling can be used to create likelihood functions of age for partially bleached OSL data, incorporating uncertainty from two sources: the sensitivity of the age model to each aliquot or grain, and the assumed width in the well-bleached population of grains. The main advantages of bootstrap likelihoods are:

- An improved assessment of uncertainty in OSL ages derived from partially bleached samples.
- The possibility of incorporating data from partially bleached OSL samples into chronological models using Bayes' theorem.

The bootstrap likelihood protocol provides a framework for attaching future improvements in OSL methods, e.g. a different age model, or better assessment of dose-rate variation between grains. Maximum benefit from this protocol will occur for sequences with high-resolution sampling and detailed stratigraphic information. This protocol is a new and promising approach that provides large benefits over presently used (non-bootstrap) methods, and we hope it will be further expanded and developed in the future. Finally we note that for our study site, single-grain OSL measurements were inefficient and added no value to the small-aliquot data.

Acknowledgements

Scripts for the bootstrap likelihoods have been written in Matlab, and can be obtained from www.ncl.tudelft.nl or directly from the authors. The age models were translated from the S-Plus versions of Arnold et al. (2009). We thank Noortje Hobo for the samples, and for discussion on dating methods. The manuscript was much improved following comments by Rex Galbraith and two further reviewers. The authors are supported by NWO/STW grant DSF.7553.

References

Arnold, L.J., Roberts, R.G., 2009. Stochastic modelling of multi-grain equivalent dose (De) distributions: Implications for OSL dating of sediment mixtures. Quaternary Geochronology 4, 204-230.

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

Bailey, R.M., Arnold, L.J., 2006. Statistical modelling of single grain quartz De distributions and an assessment of procedures for estimating burial dose. Quaternary Science Reviews 25, 2475-2502.

Ballarini, M., Wallinga, J., Duller, G.A.T., Brouwer, J.C., Bos, A.J.J., Van Eijk, C.W.E. 2005. Optimizing detection filters for single-grain optical dating of quartz. Radiation Measurements 40, 5-12.

Benito, G., Thorndycraft, V.R., Rico., M., Sanchez-Moya, Y., Sopena, A., 2008. Palaeoflood and floodplain records from Spain: Evidence for long-term climate variability and environmental changes. Geomorphology 101, 68-77.

Blockley, S.P.E., Blaauw, M., Bronk Ramsey, C., van der Plicht, J., 2007. Building and testing age models for radiocarbon dates in Lateglacial and Early Holocene sediments. Quaternary Science Reviews 26, 1915-1926.

Bøtter-Jensen, L., Bular, E., Duller, G.A.T., Murray, A.S., 2000. Advances in luminescence instrument systems. Radiation Measurements 41, 1020-1025.

Bronk Ramsey, C., 1995. Radiocarbon calibration and analysis of stratigraphy: the OxCal program. Radiocarbon 37, 425-430.

Bronk Ramsey, C., 2008. Deposition models for chronological records. Quaternary Science Reviews 27, 42-60.

Bronk Ramsey, C. 2009. Bayesian analysis of radiocarbon dates. Radiocarbon, 51(1), 337-360.

Bronk Ramsey, C., Dee, M.W., Rowland, J.M., Higham, T.F.G., Harris, S.A., Brock, F., Quiles, A., Wild, E.M., Marcus, E.S., Shortland, A.J., 2010. Radiocarbon-based chronology for dynastic Egypt. Science, 328, 1554-1557.

Buck, C., Kenworthy, J., Litton, C., Smith, A., 1991. Combining archaeological and radiocarbon information – a Bayesian approach to calibration. Antiquity 65, 808-821.

Busschers F.S., Kasse, C.,van Balen, R.T., Vandenberghe, J., Cohen, K.M., Weerts, H.J.T., Wallinga, J., Johns, C., Cleveringa, P., Bunnik, F.P.M., 2007. Late Pleistocene evolution of the Rhine-Meuse system in the Southern North Sea basin: imprints of climate change, sea-level oscillation and glacio-isostacy. Quaternary Science Reviews 26, 3216-3248.

Busschers, F.S., Van Balen, R.T., Cohen, K.M., Kasse, C., Weerts, H.J.T., Wallinga, J., Bunnik, F.P.M., 2008. Response of the Rhine-Meuse fluvial system to Saalian ice-sheet dynamics. Boreas 37, 377-398.

Cunningham, A.C., Wallinga, J., 2010. Selection of integration time intervals for quartz OSL decay curves. Quaternary Geochronology 5, 657-666.

Cunningham, A.C., Bakker, M.A.J., Van Heteren, S., Van Der Valk, B., Van Der Spek, A.J.F., Schaart, D.R., Wallinga, J., 2011a. Extracting storm-surge data from coastal dunes for improved assessment of flood risk. Geology 39, 1063-1066.

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

Davison, A.C., Hinkley, D.V., Worton, B.J., 1992. Bootstrap likelihoods. Biometrika 79, 113-130.

Davison, A.C., Hinkley, D.V., Worton, B.J., 1995. Accurate and efficient construction of bootstrap likelihoods. Statistics and Computing 5, 257-264.

Duller, G.A.T., 2008. Single-grain optical dating of Quaternary sediments: why aliquot size matters in luminescence dating. Boreas 37, 589-612.

Efron, B. 1979., Bootstrap methods: another look at the Jackknife. Annals of Statistics 7, 1-26.

Efron, B., Tibshirani, R.J., 1993. An introduction to the Bootstrap. Chapman & Hall, New York.

Fitzsimmons, K., 2011. An assessment of the luminescence sensitivity of Australian quartz with respect to sediment history. Geochronometria 33, 199-208.

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

Galbraith, R.F., Roberts, R.G., Laslett, G.M., Yoshida, H., Olley, J.M., 1999. Optical dating of single and multiple grains of quartz from jinmium rock shelter, northern Australia, part 1, Experimental design and statistical models. Archaeometry 41, 339-364.

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

Hobo, N., Makaske, B., Middelkoop, H., Wallinga, J., 2010. Reconstruction of floodplain sedimentation rates: a combination of methods to optimize estimates. Earth Surface Processes and Landforms 35, 1499-1515.

Jacobi, R.M., Higham, T.F.G., 2009. The early Lateglacial re-colonization of Britain: new radiocarbon evidence from Gough's Cave, southwest England. Quaternary Science Reviews 28, 1895-1913.

Jain, M., Murray, A. S., Bøtter-Jensen, L., 2004. Optically stimulated luminescence dating: How significant is incomplete light exposure in fluvial environments? Quaternaire 15, 143-157.

Lukas, S., Spencer, J.Q.G., Robinson, R.A.J., Benn, D.I., 2007. Problems associated with luminescence dating of Late Quaternary glacial sediments in the NW Scottish Highlands. Quaternary Geochronology 2, 243-248.

Mayya, Y.S., Morthekai, P., Murari, M.K., Singhvi, A.K., 2006. Towards quantifying beta microdosimetric effects in single-grain quartz dose distribution. Radiation Measurements 32, 57-73.

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

Newton, M.A., Geyer, C.T., 1994. Bootstrap recycling: a Monte Carlo alternative to the nested bootstrap. Journal of the American Statistical Association 89, 905-912.

Owens, P.N., Walling, D.E., Leeks, G.J.L., 1999. Use of floodplain sediment cores to investigate recent historical changes in overbank sedimentation rates and sediment sources in the catchment of the River Ouse, Yorkshire, UK. Catena 36, 21–47.

Parnell. A.C., Buck, C.E., Doan, T.K, 2011. A review of statistical chronology models for high-resolution, proxy-based Holocene palaeoenvironmental reconstruction. Quaternary Science Reviews 30, 2948-2960.

Rhodes, E.J., Bronk Ramsey, C., Outram, Z., Batt, C., Willis, L., Dockrill, S., Bond, J., 2003. Bayesian Methods applied to the interpretation of multiple OSL dates: high precision sediment ages from Old Scatness Broch excavations, Shetland Isles. Quaternary Science Reviews 22, 1231-1244.

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

Rittenour, T. M., Goble, R. J. & Blum, M.D., 2005. Development of an OSL chronology for late Pleistocene channel belts in the lower Mississippi valley. Quaternary Science Reviews 24, 2539-2554.

Rodnight, H., Duller, G.A.T., Tooth, S., Wintle, A.G., 2005. Optical dating of a scroll-bar sequence on the Klip River, South Africa, to derive the lateral migration rate of a meander bend. The Holocene 15, 802-811.

Rodnight, H., Duller, G.A.T., Wintle, A.G., Tooth, S., 2006. Assessing the reproducibility and accuracy of optical dating of fluvial deposits. Quaternary Geochronology 1, 109-120.

Thrasher, I.M., Mauz, B., Chiverrell, R.C., Lang, A., Thomas, G.S.P., 2009, Testing an approach to OSL dating of Late Devensian glaciofluvial sediments of the British Isles. Journal of Quaternary Science 24, 785-801.

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

Thomsen, K.J., Murray, A.S., Bøtter-Jensen, L., Kinahan, J., 2007. Determination of burial dose in incompletely bleached fluvial samples using single grains of quartz. Radiation Measurements 42, 370-379.

Wallinga, J., 2002. Optically stimulated luminescence dating if fluvial deposits: a review. Boreas 31, 303-322.

Wallinga, J., Hobo, N., Cunningham, A.C., Versendaal, A.J., Makaske, B., Middelkoop, H., 2010. Sedimentation rates on embanked floodplains determined through quartz optical dating. Quaternary Geochronology 5, 170-175.

Figure Captions

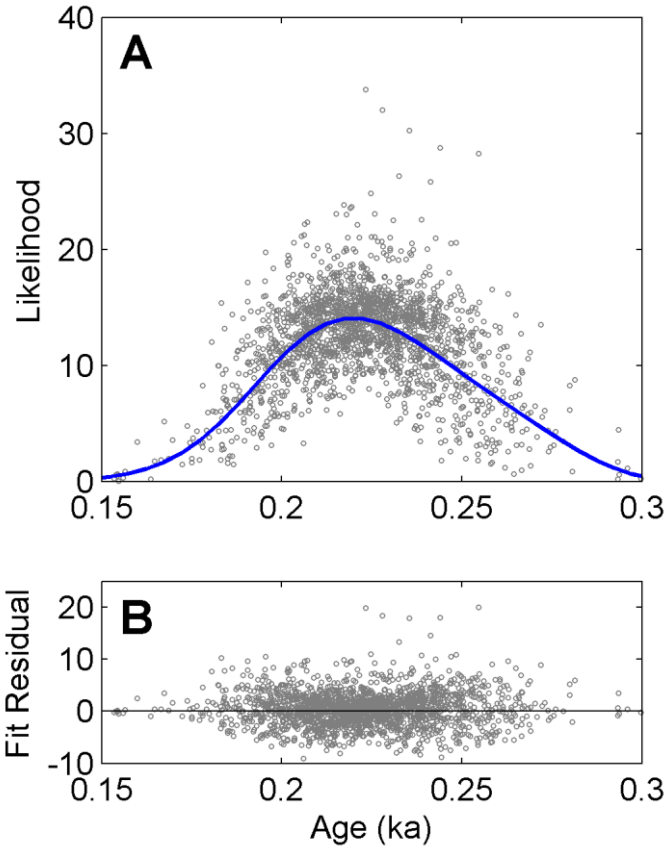

Fig. 1. Construction of the bootstrap likelihood. (a) 2000 bootstrap replicates of the minimum age have been assigned a likelihood value using bootstrap recycling. The data is fitted with a 6-degree polynomial to estimate the likelihood as a function of age. (b) Fitting residuals. The curve was fitted on the logged data.

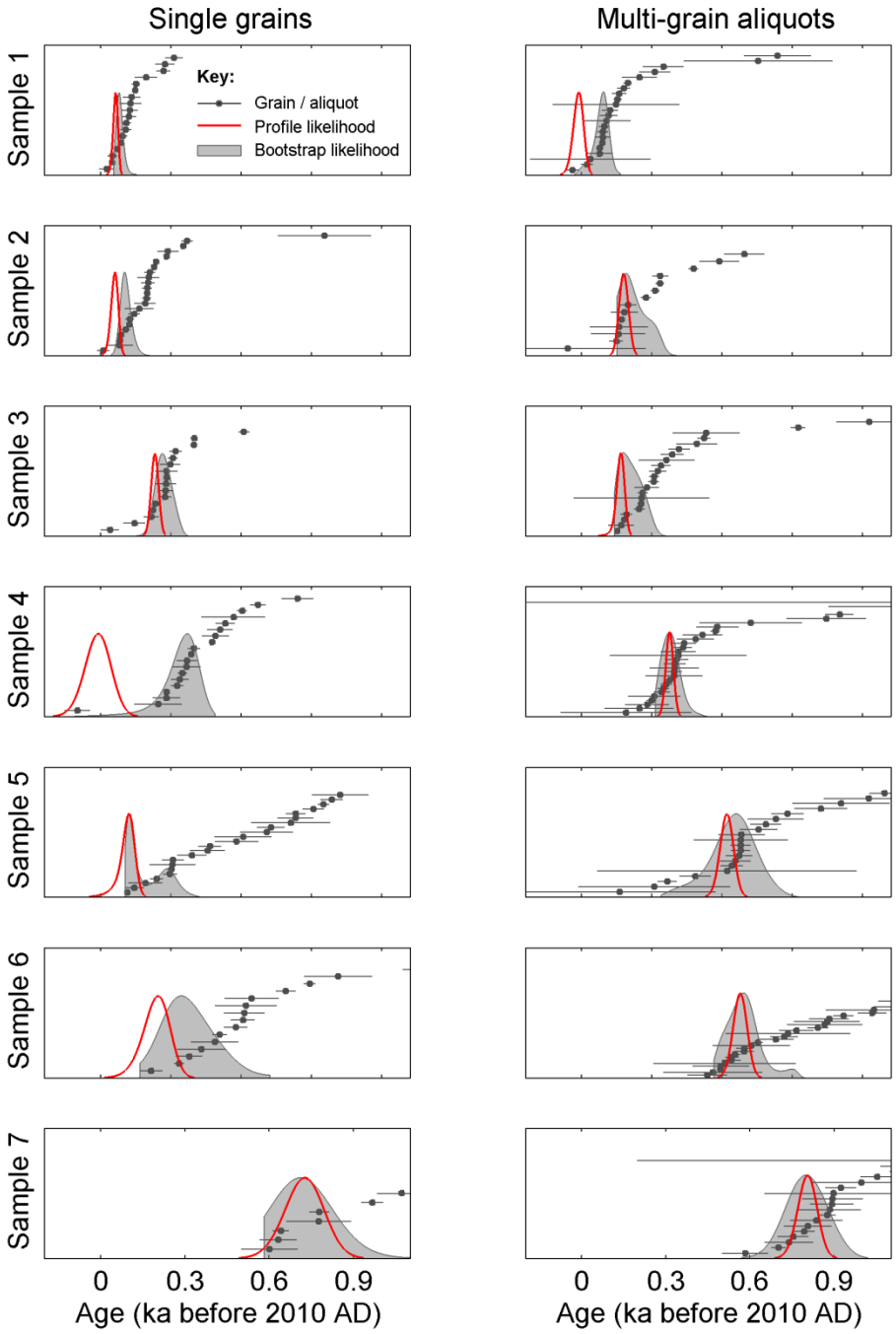

Fig. 2. Bootstrap likelihoods for a sequence of fluvial samples, using single-grain (left) and small-aliquot data (right). The samples come from a single core, and are plotted in stratigraphic order. Also plotted is the $MAM3_{ul}$ profile likelihood for each sample, which would ordinarily provide the confidence intervals, and the D_e for each

accepted aliquot or grain. The likelihoods are normalised by height. The x-axis has been converted to age using the sample-specific dose rates.

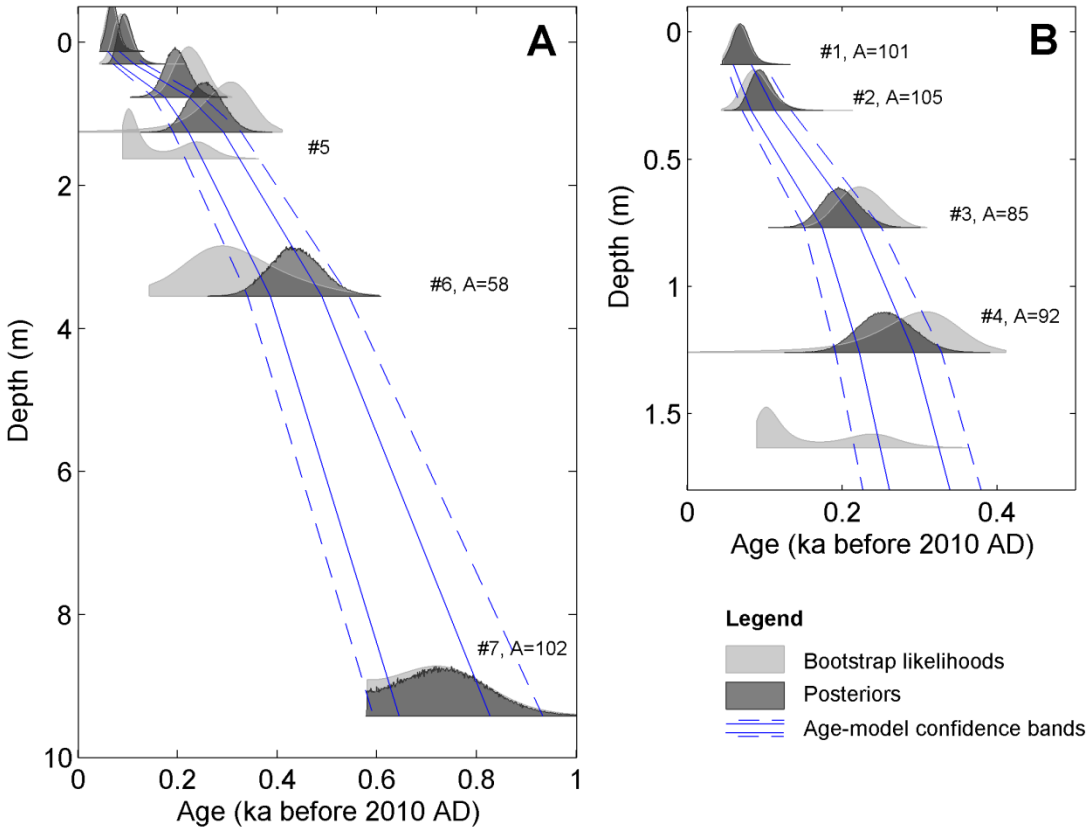

Fig. 3. (a) Age-depth model for a sequence of fluvial samples using **single grains** of quartz. (b) enlargement of the upper part of the sequence. Bootstrap likelihoods were created using the procedure described in section 2.3. The likelihoods were combined with prior information using OxCal 4.1; model specifications are given in section 3. Sample 5 was omitted from the final OxCal model due to a poor agreement score. Age model 68% and 95% confidence regions are shown, using linear interpolation

between the posteriors. The sample number and agreement index for each sample is stated beside the curves.

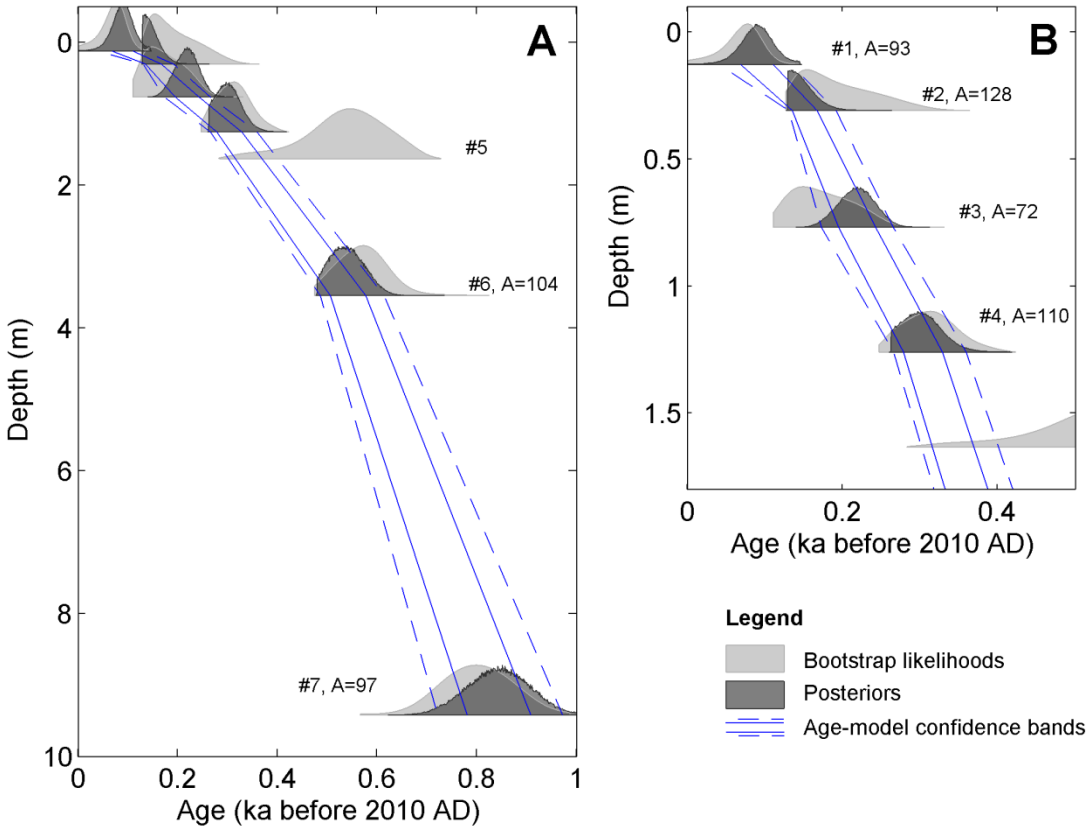

Fig. 4. (a) Age-depth model for a sequence of fluvial samples using **small-aliquots** (2-3 mm) of quartz. (b) enlargement of the upper part of the sequence. Bootstrap likelihoods were created using the procedure described in section 2.3. The likelihoods were combined with prior information using OxCal 4.1; model specifications are given in section 3. Sample 5 was omitted from the final OxCal model due to a poor agreement score. Age model 68% and 95% confidence regions are shown, using linear interpolation between the posteriors. The sample number and agreement index for each sample is stated beside the curves.

Sample No.	Lab code	Depth (m)	Overdispersion	
			Single grains	Small aliquots
1	NCL-1107140	0.13	0.50 ± 0.10	0.73 ± 0.14
2	NCL-1107141	0.31	0.52 ± 0.09	0.79 ± 0.11
3	NCL-1107142	0.77	0.89 ± 0.15	0.58 ± 0.08
4	NCL-1107143	1.26	0.73 ± 0.16	0.54 ± 0.08
5	NCL-1107144	1.64	0.81 ± 0.14	0.76 ± 0.10
6	NCL-1107146	3.55	0.56 ± 0.11	0.69 ± 0.07
7	NCL-1107147	9.42	0.63 ± 0.11	0.37 ± 0.05

Table 1. Overdispersion in the equivalent-dose data for each sample, calculated using the central-age model (Galbraith et al., 1999).