# Making (non)sense of age models: a guide for (ab)users

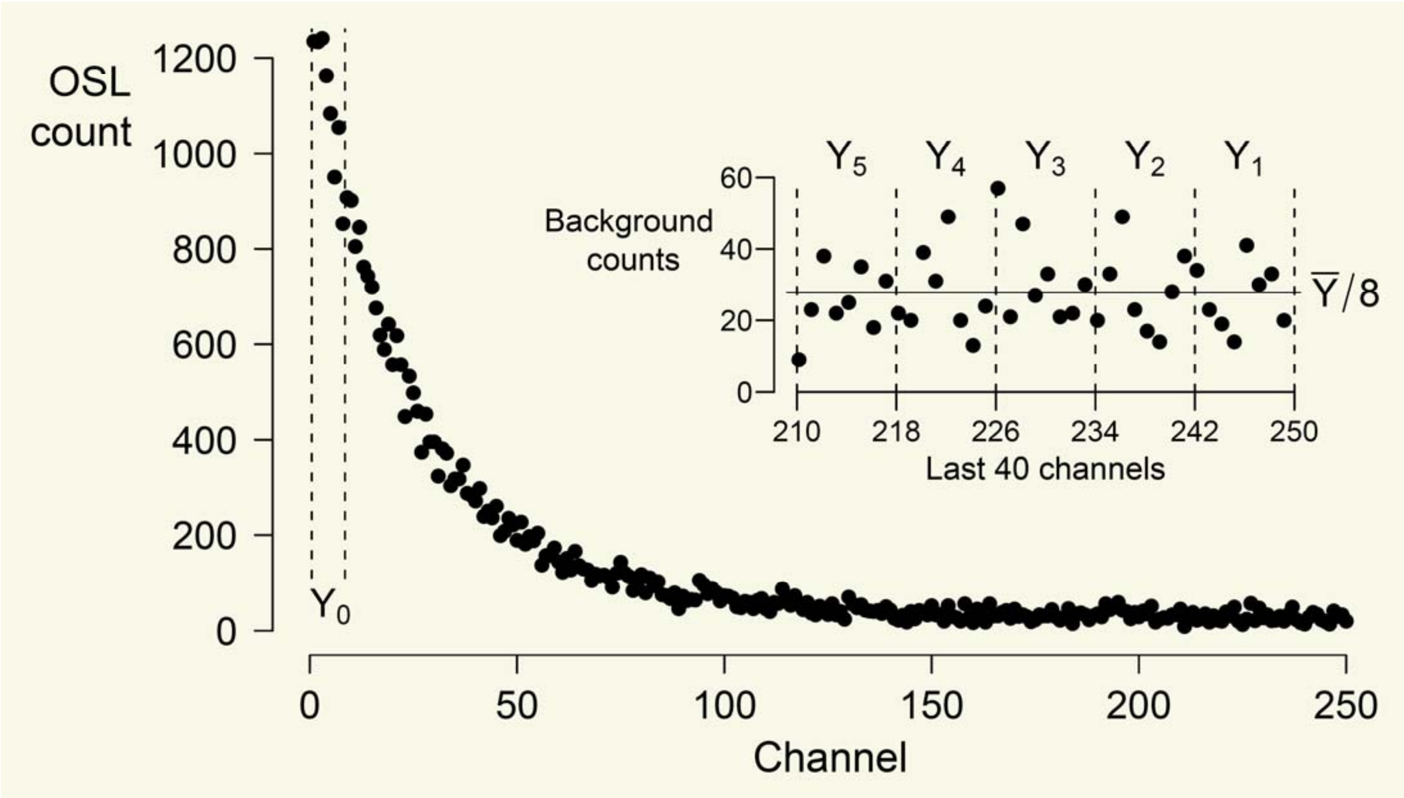
**Richard 'Bert' Roberts**

Centre for Archaeological Science,
School of Earth & Environmental Sciences,
University of Wollongong, Australia

# Talk outline

- ‘Age’ models for OSL data – why bother?
- Estimating your errors without erring
- Displaying your data simply and correctly
- Data behaving badly – ‘over-dispersion’
- Age models – context is everything!
- Age models – some practical tips for beginners

# Photon counts from optical stimulation of a quartz grain




OSL counts in $N$ = 250 consecutive channels (inset: magnified graph of counts in the last 40 channels). $Y_0$ = the total count over the first $n$ channels, and $Y_1$, …, $Y_k$ and $\overline{Y}$ are used to estimate the background count. Here, $n$ = 8 and $k$ = 5.

$$rse(\hat{\mu}_s) = \frac{\sqrt{Y_0 + \overline{Y} / k}}{Y_0 - \overline{Y}}$$

$\hat{\mu}_s$ = estimate of the background-corrected ('net') OSL signal, <u>assuming</u> the background counts have a Poisson distribution.

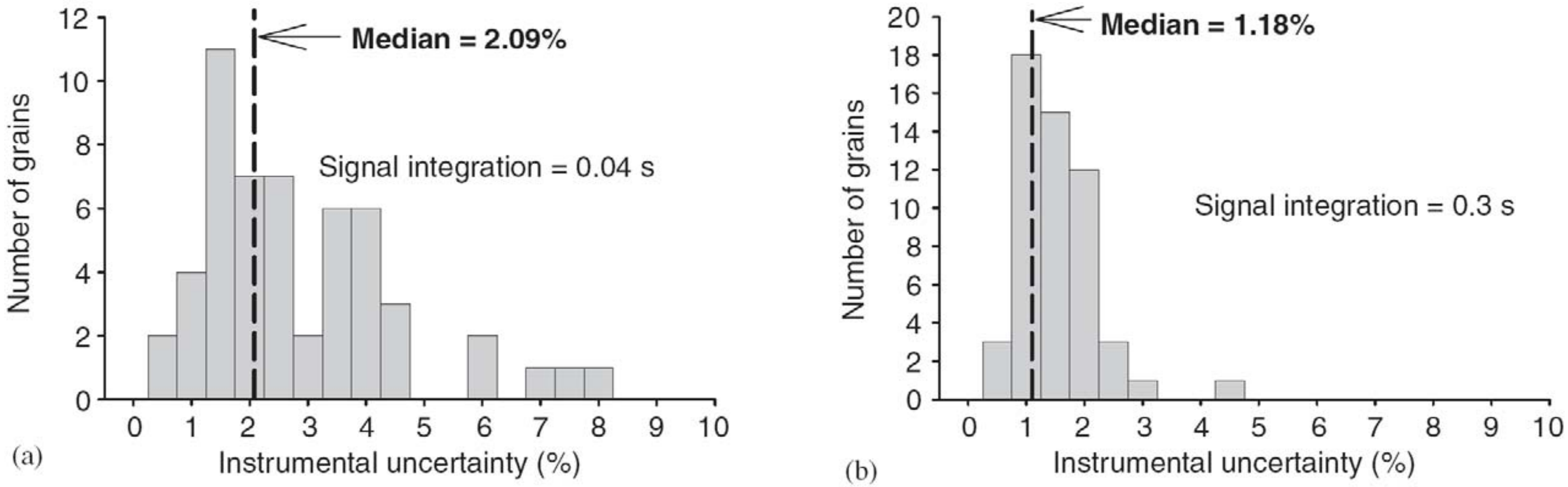
## But do they?

- Calculate the ratio of the variance (i.e., the square of the standard deviation) to the mean of the *k* background counts
- Variance/mean ratio = 1 if Poisson, and >1 if 'extra-Poisson'
- For data with an extra-Poisson component, use equations given by Rex Galbraith (*Ancient TL* 2002):

$$rse(\hat{\mu}_s) \approx \frac{\sqrt{Y_0 + \overline{Y} / k}}{Y_0 - \overline{Y}} \times \sqrt{1 + \frac{\hat{\sigma}^2}{\overline{Y}}} \quad \text{where} \quad \hat{\sigma}^2 = \left( \frac{1}{k-1} \sum_{j=1}^{k} (Y_j - \overline{Y})^2 \right) - \overline{Y}$$

# Other errors to include

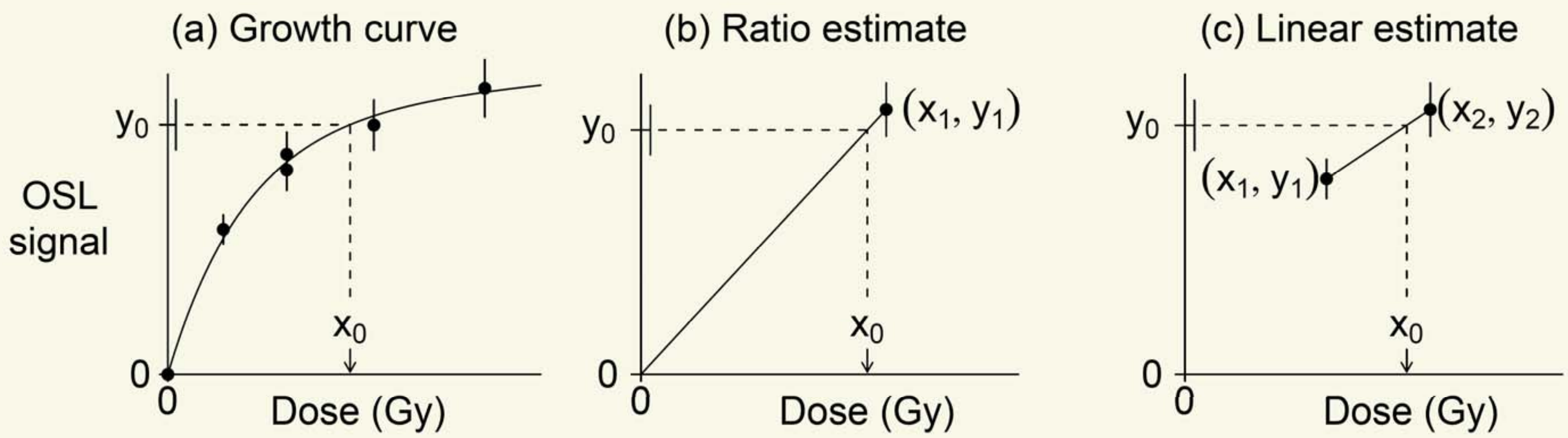
- Instrumental (ir)reproducibility
  - Measure for <u>each</u> instrument
  - Determine for <u>specific</u> signal-integration periods
  - Simple method described by Jacobs *et al*. (*Radiat Meas* 2006)



# Estimation of the equivalent dose, $D_e$

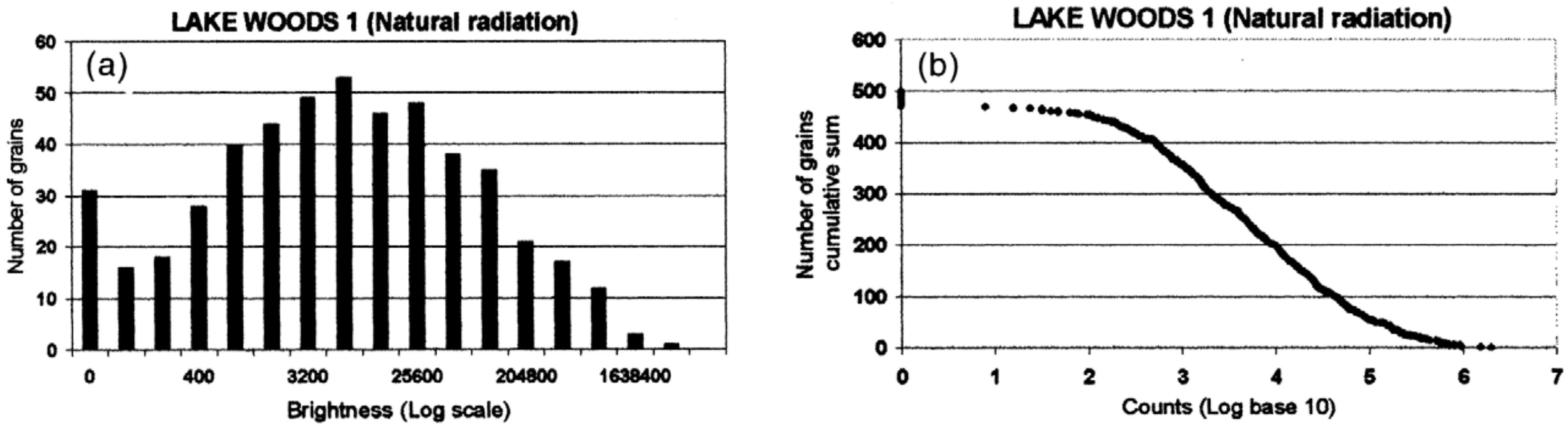
The data are OSL signals $y_1, y_2, \ldots, y_n$ (and their precisions) obtained from given known doses $x_1, x_2, \ldots, x_n$ (corrected for background and sensitivity change) plus the natural OSL signal $y_0$.

We want to estimate the natural dose $x_0$ that produced $y_0$ (and its precision).



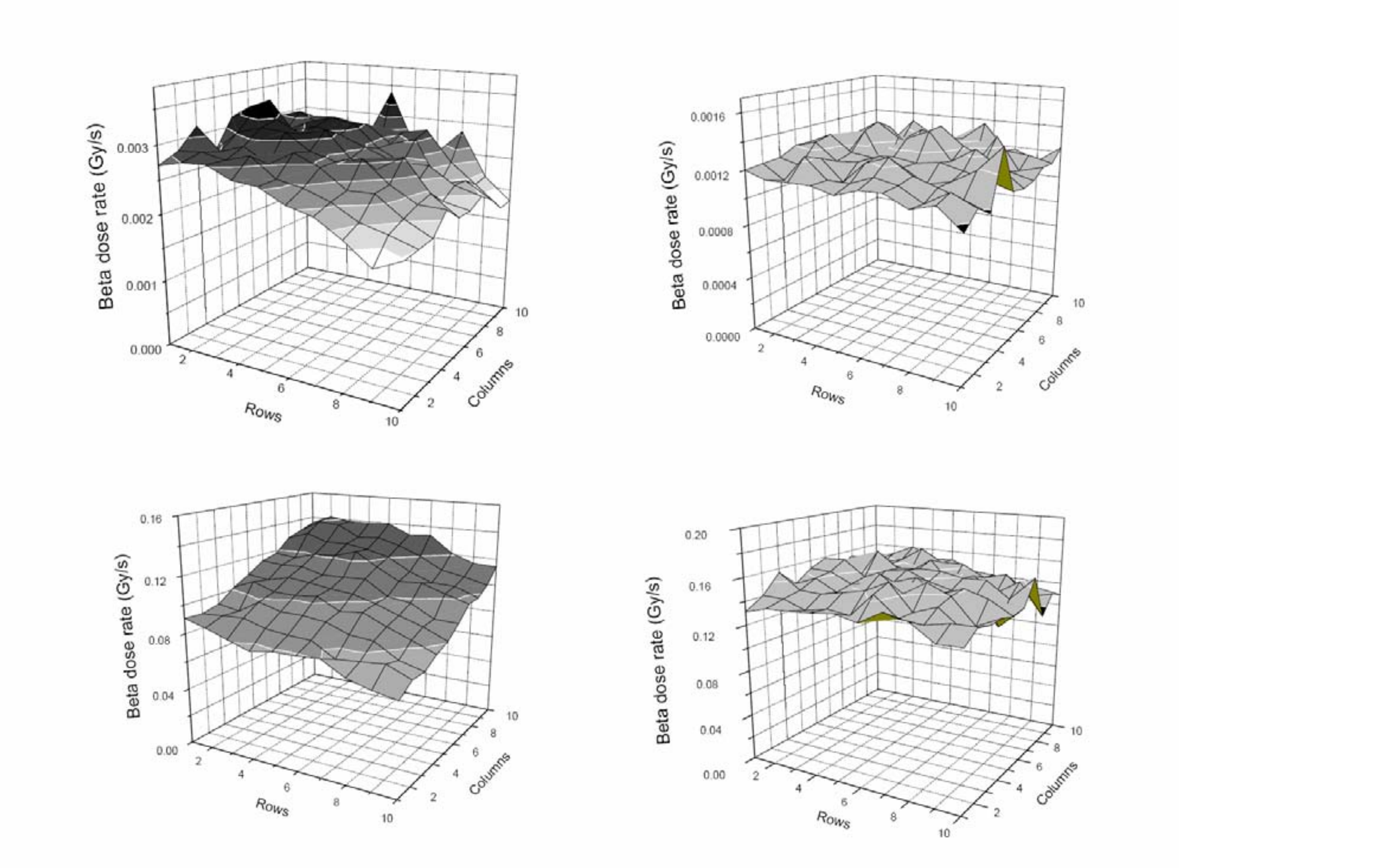
# Curve-fitting errors

- Determine by Monte Carlo simulation
  - What form is the frequency distribution of OSL intensities at each dose point – normal (Gaussian) or log-normal?
  - Log-normal distributions of single-grain OSL intensities reported by McCoy *et al*. (*Radiat Meas* 2000)
  - Appropriate MC methods described by Yoshida *et al*. (*Radiat Meas* 2003) and Duller (*Ancient TL* 2007)

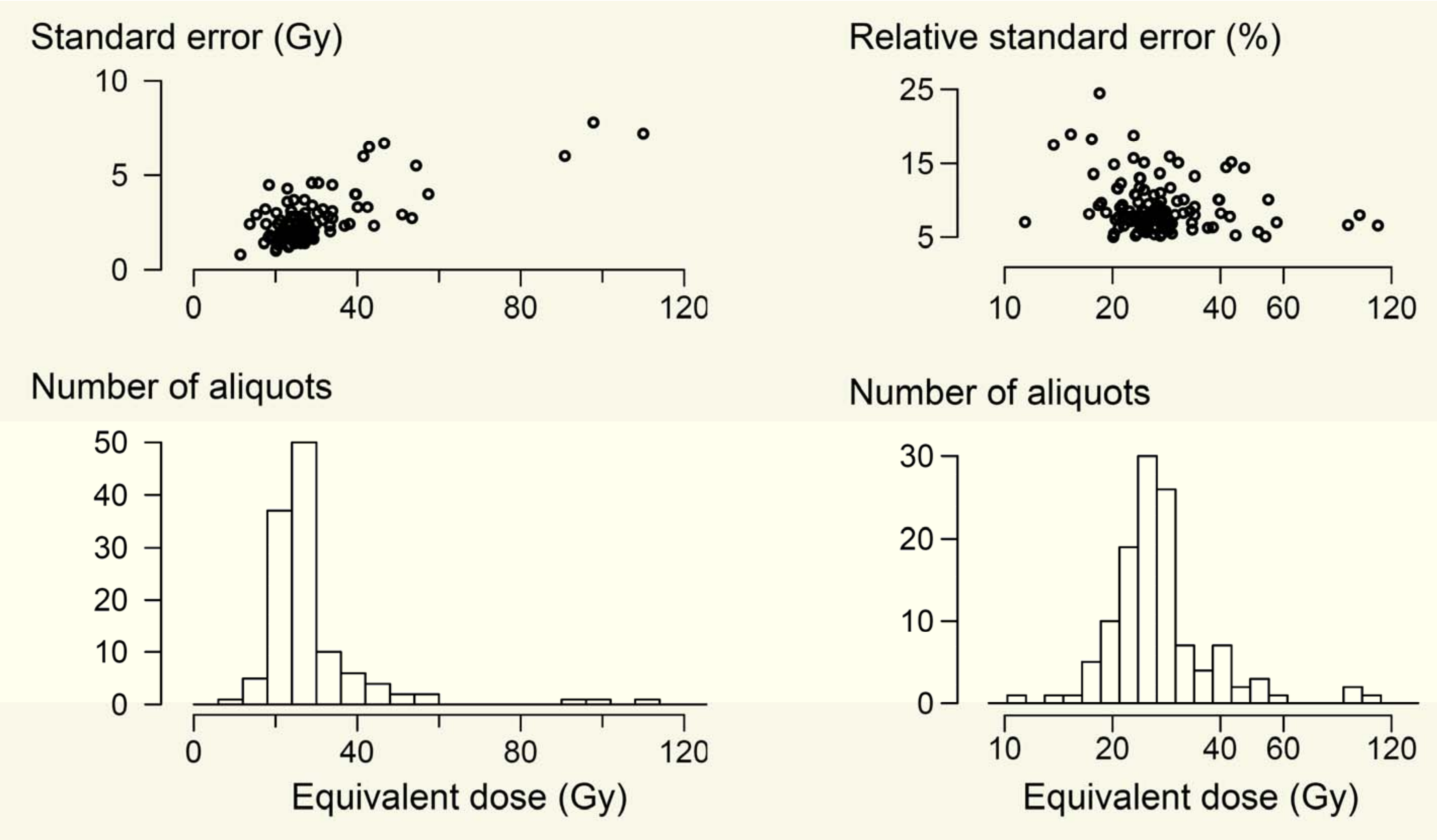


# Beta-source variability to single grains

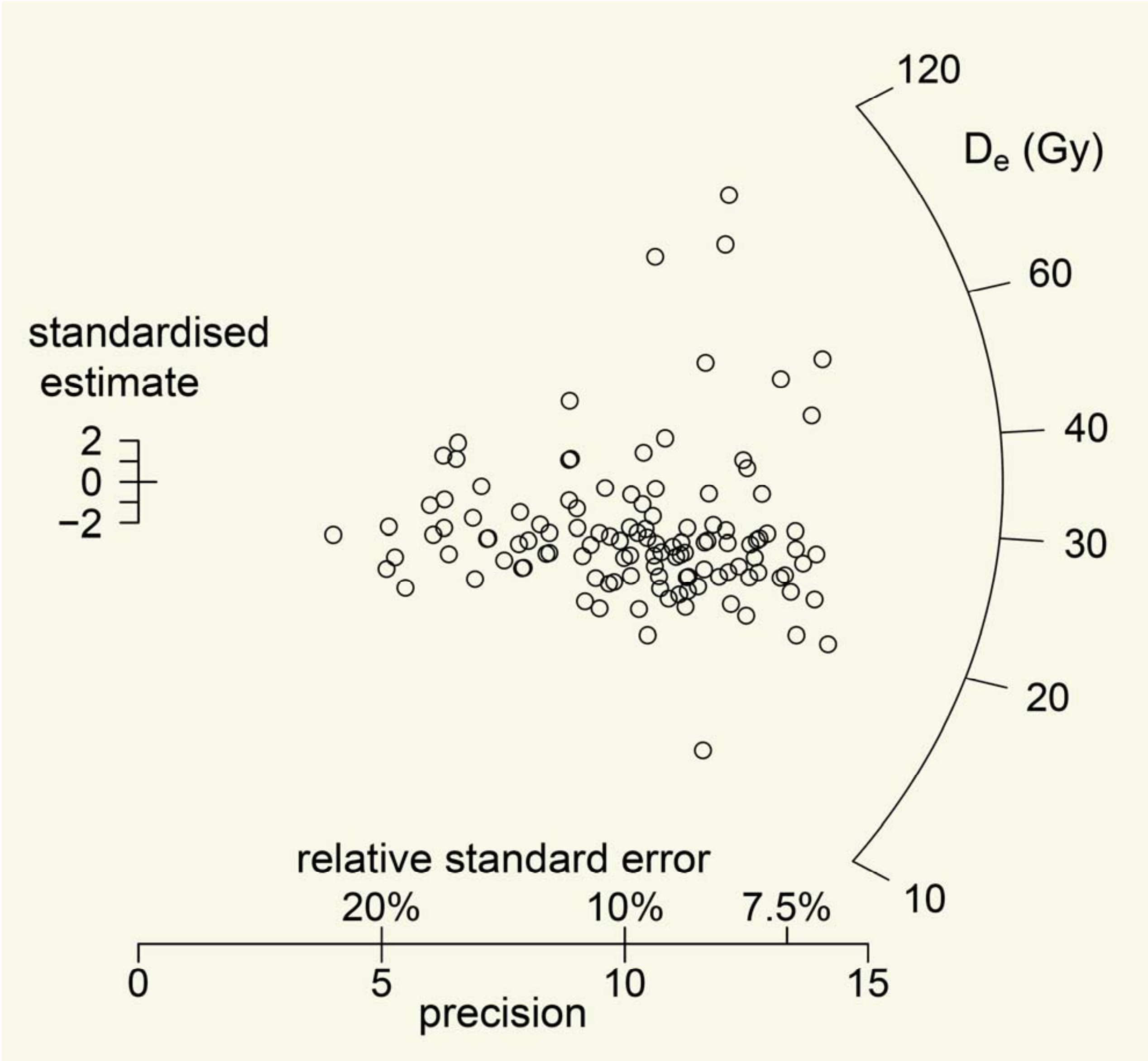
- Cross-disc spatial variation in dose rate from laboratory β sources
- Measure for <u>each</u> instrument, e.g. Ballarini *et al*. (*Ancient TL* 2006)



# $D_e$ estimates for 120 aliquots



# Radial plot of $D_e$ estimates for 120 aliquots



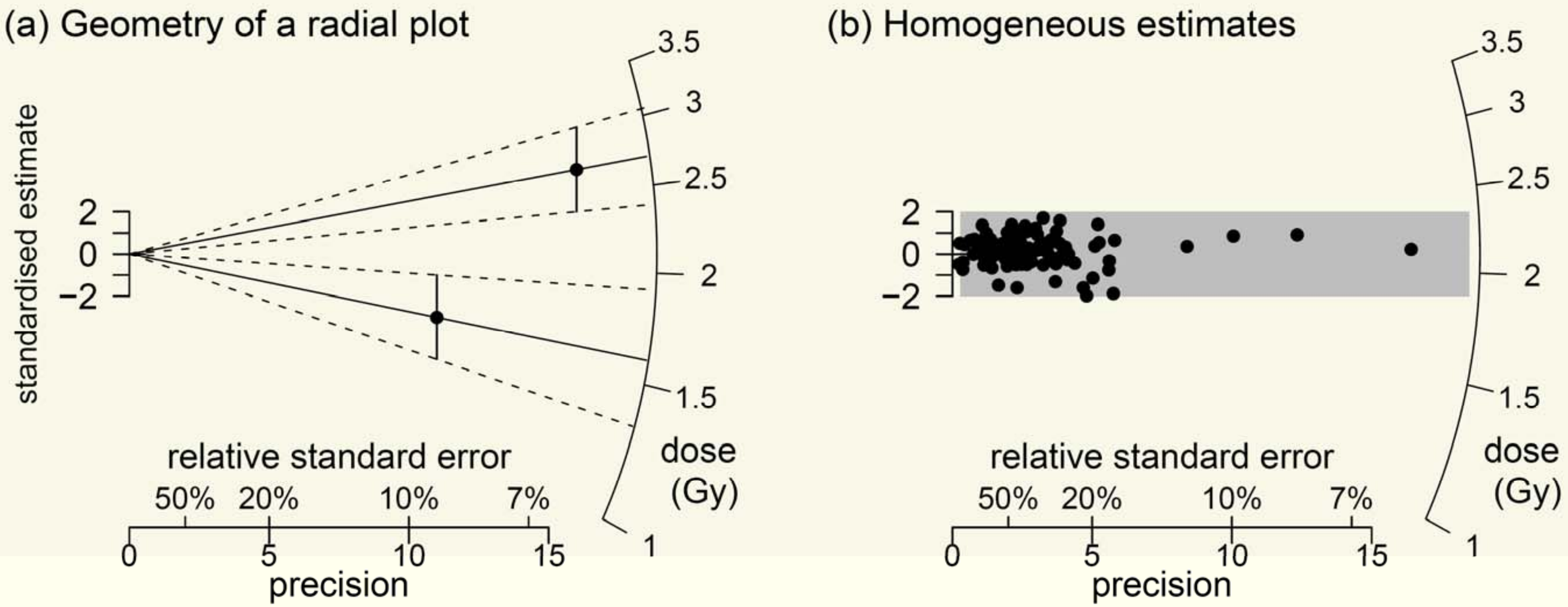
# Radial plots

Data are estimates $z_1, z_2, \ldots, z_n$ with standard errors $\sigma_1, \sigma_2, \ldots, \sigma_n$

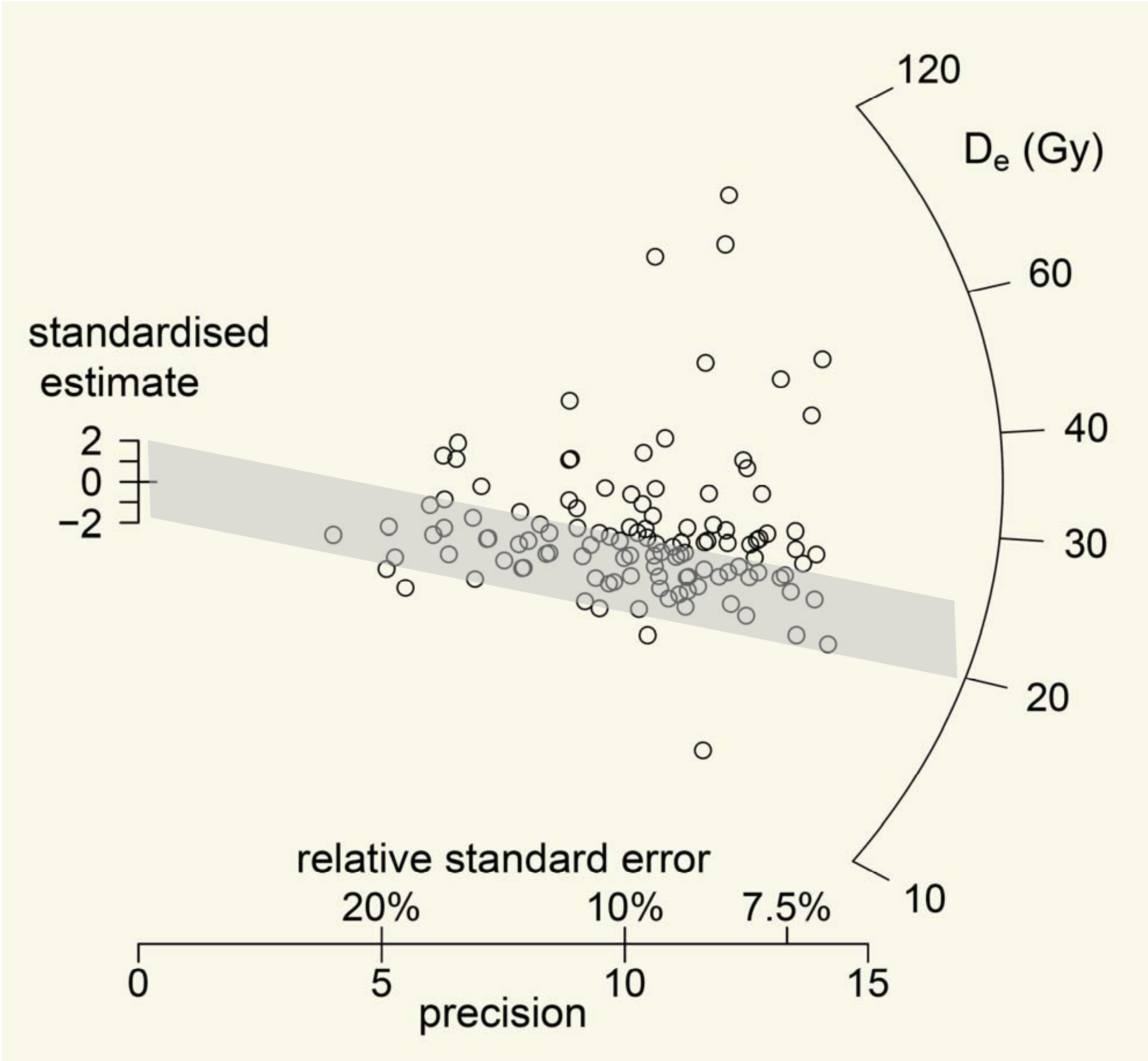
Draw a scatter plot of $y_i = (z_i - z_0)/\sigma_i$ against $x_i = 1/\sigma_i$

Then $z_i - z_0 = y_i/x_i =$ slope of line through (0,0) and $(x_i, y_i)$

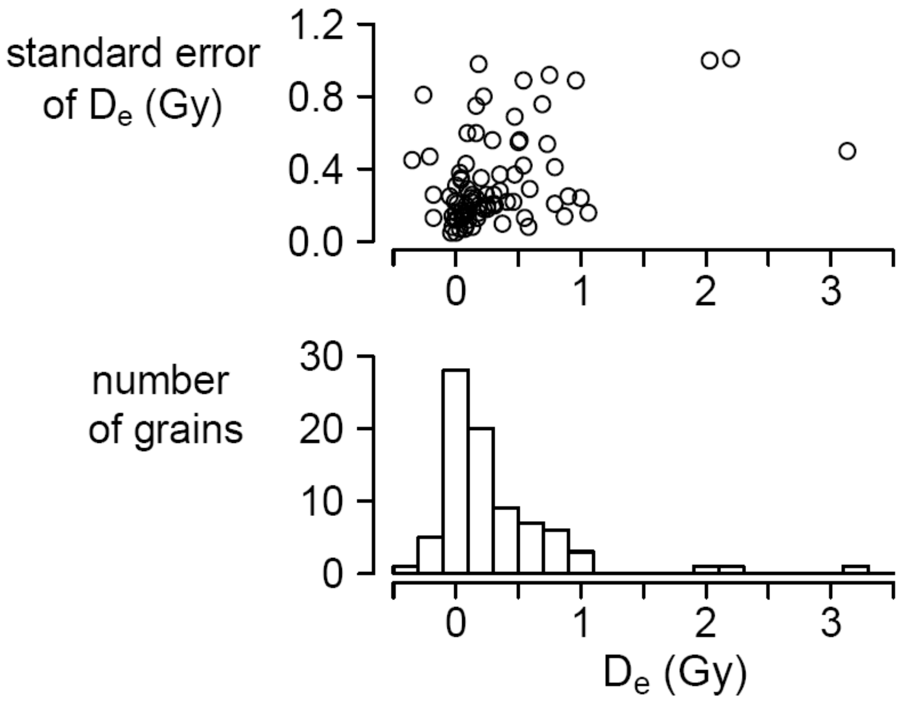
and each $y_i$ has unit standard deviation

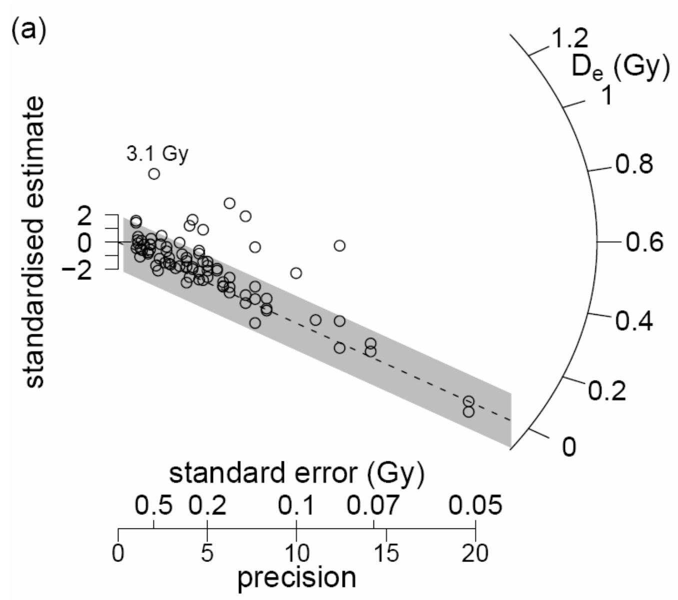


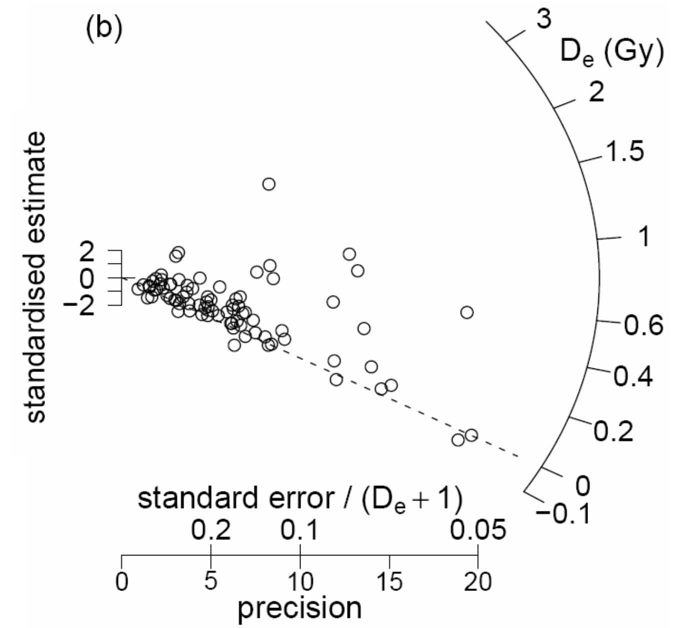
# Radial plot of $D_e$ estimates for 120 aliquots



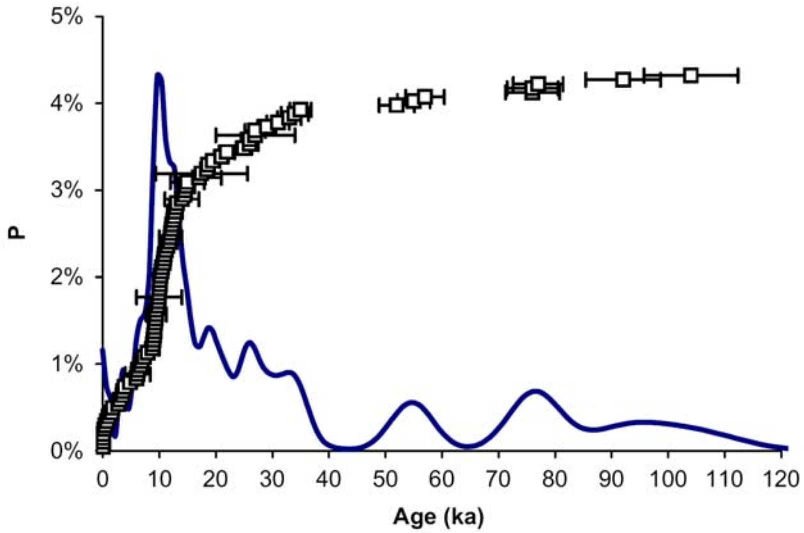
# $D_e$ estimates for young and modern samples



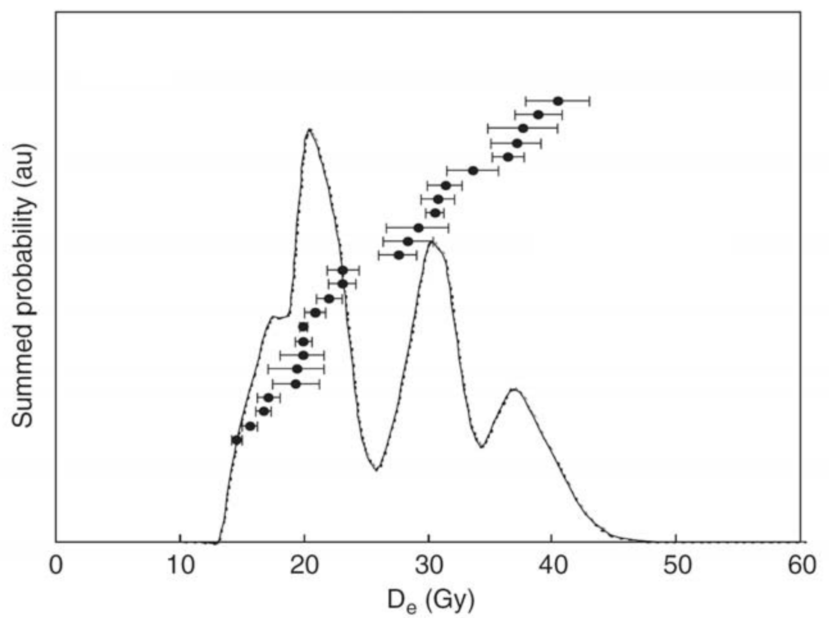



# Weighted histograms (probability density plots)

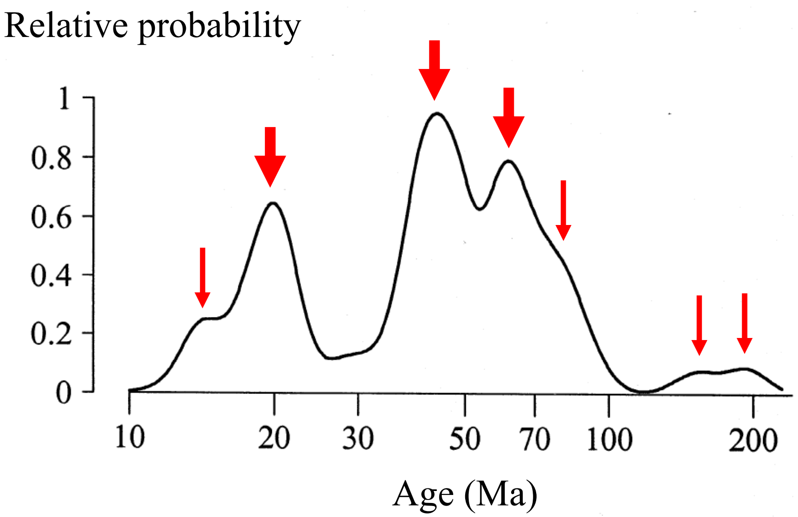


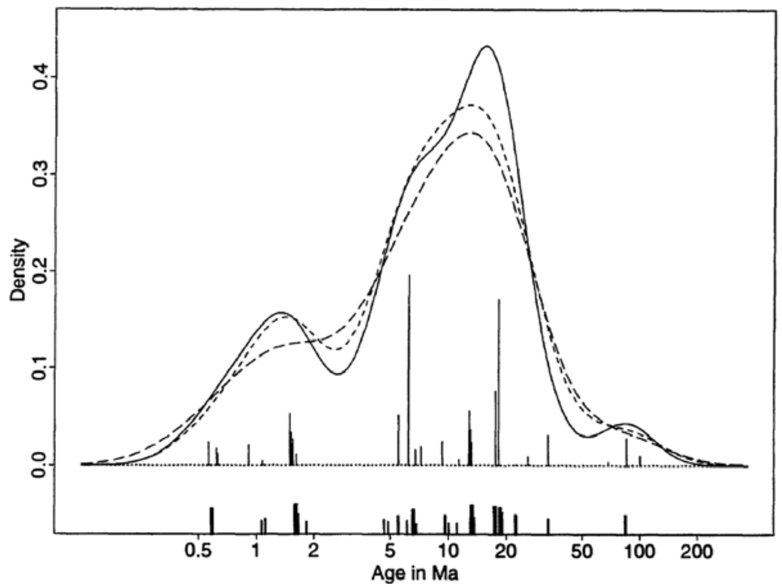

dune sediments

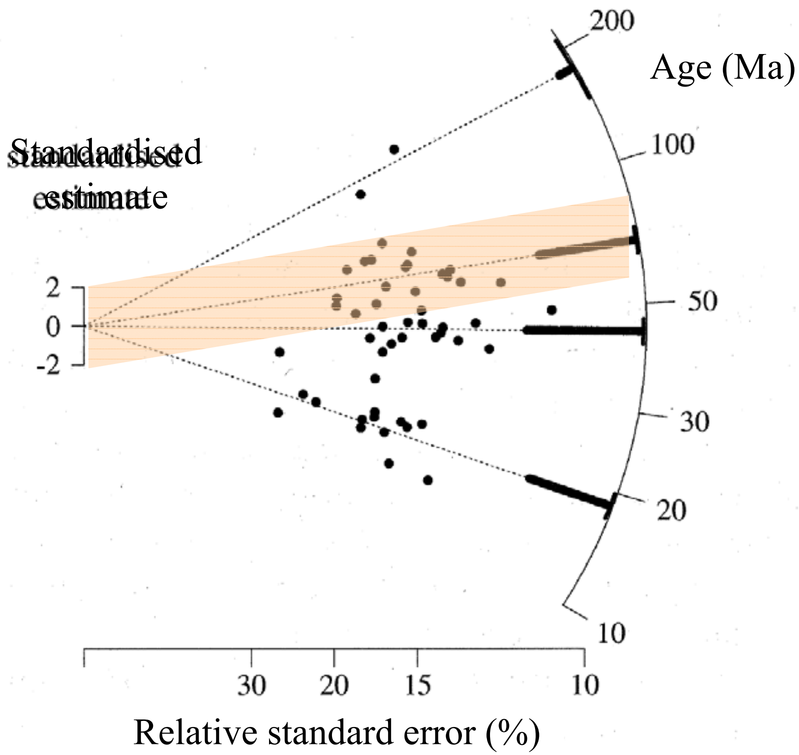
river sediments



# Probability density plots: conflating 2 sources of error

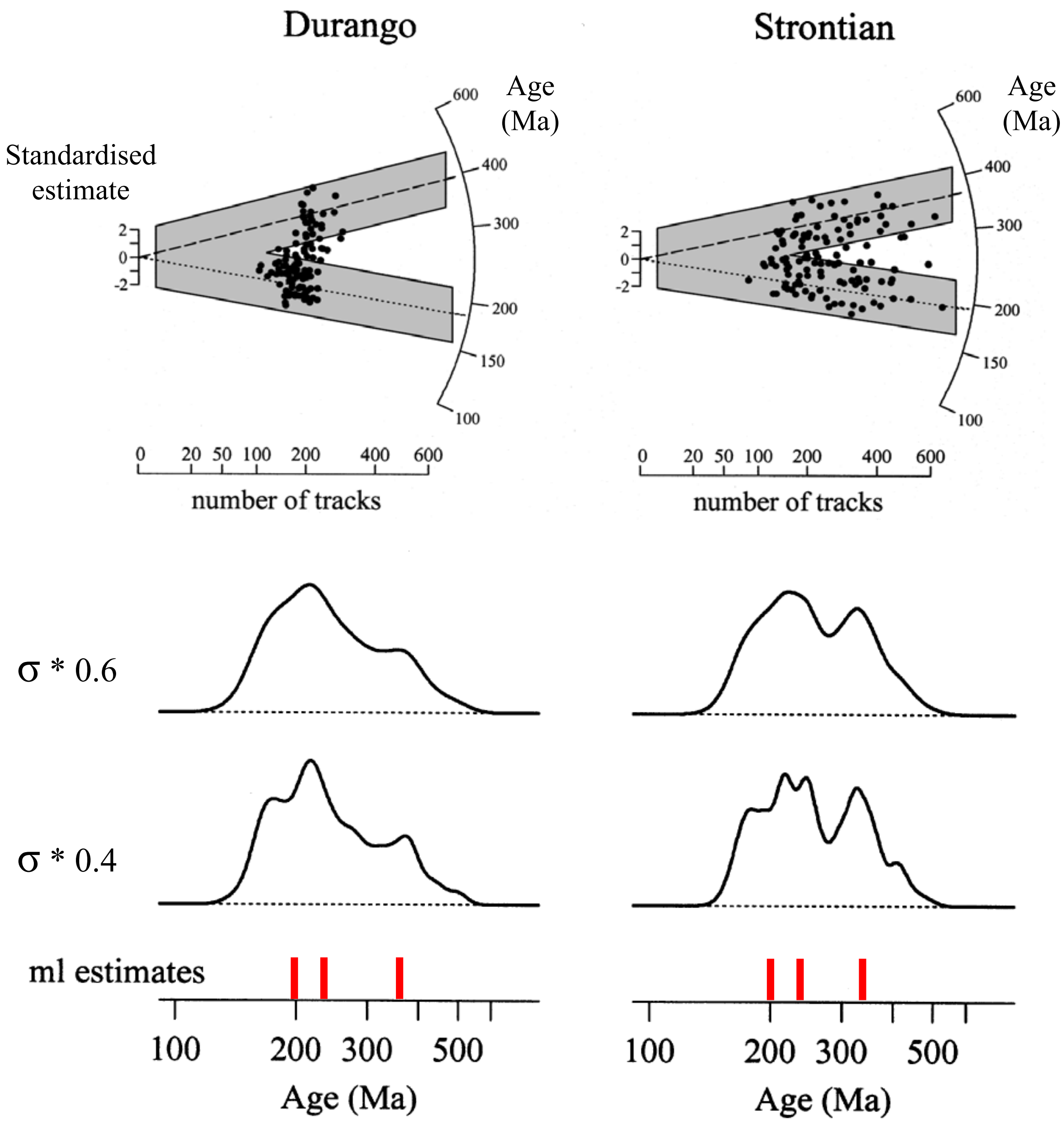



Galbraith (*Radiat Meas* 1998)

Goutis (*J Am Stat Assoc* 1997)



Durango
Strontian
Standardised estimate
Age (Ma)
2
0
-2
600
400
300
200
150
100
0 20 50 100 200 400 600
number of tracks
σ * 0.6
σ * 0.4
ml estimates
100
200 300 500
Age (Ma)

# Over-dispersion

- The extra spread in $D_e$ values remaining after having taken into account measurement uncertainties:
    - photon counting statistics
    - background correction (extra-Poisson?)
    - instrumental (ir)reproducibility
    - curve-fitting errors
    - beta-source heterogeneity (for single grains)
- Relative standard deviation of the $D_e$ estimates above and beyond all known uncertainties, expressed in %
- Denoted by $\sigma_b$ by Galbraith *et al*. (*Radiat Meas* 2005)
- Ubiquitous in natural samples and lab-dosed samples
    - dose-recovery tests
    - recycling ratios

# Why do we get over-dispersion?

- Because of ‘experimental’ and ‘natural’ variation
  - experimental errors are reducible
  - but natural variation is <u>inherent</u>

- Variation between aliquots/grains in the lab
  - OSL source traps not filled to same extent
    - before burial: partial bleaching
    - after burial: sediment mixing, micro-dosimetry variations
  - heating/bleaching don’t empty all traps uniformly
  - differences in thermal transfer effects

- Non-identical field and lab conditions
  - bleaching spectra
  - types of ionising radiation
  - dose rates
  - defect migration with time/temperature

# How much over-dispersion do we get?

- Factorial experiment (Galbraith *et al*., *Radiat Meas* 2005)
- Six replicates of a $2^3$ factorial design. The 3 factors:

1. size of test dose – 0.5 or 5 Gy
2. preheat temperature – 180°C or 260°C
3. size of aliquot – 'small' (8 grains) or 'large' (80 grains)

| | |
|---|---|
| S1 | K166, annealed then given a laboratory gamma dose of 2.74 Gy |
| S2 | K162, bleached then given a laboratory beta dose of 46 Gy |
| S3 | AC150, bleached then given a laboratory beta dose of 25 Gy |
| S4 | AC150, natural (field) dose of $\sim$ 25 Gy |
| S5 | K162, bleached then given a laboratory beta dose of 2.74 Gy |
| S6 | K162, natural (field) dose of $\sim$ 46 Gy |

| Sample and dose | Aliquot size | Recycling ratio | | | | Natural and surrogate natural | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\sigma_b$ | (95% CI) | $\sigma_w$ | $\sigma_e$ | $\sigma_b$ | (95% CI) | $\sigma_w$ | $\sigma_e$ |
| S1: K166, $\gamma$ dose 2.74 Gy | Small | 1.3 | (0.8–2.0) | 2.9 | 3.2 | 4.7 | (3.5–6.5) | 2.7 | 5.4 |
| | Large | 0.9 | (0.6–1.3) | 0.4 | 1.0 | 0.8 | (0.6–1.2) | 0.4 | 0.9 |
| S2: K162, $\beta$ dose 46 Gy | Small | 1.5 | (0.9–2.5) | 2.5 | 2.9 | 1.3 | (0.5–2.5) | 2.4 | 2.7 |
| | Large | 1.0 | (0.6–1.8) | 2.4 | 2.6 | 1.5 | (0.9–2.5) | 2.4 | 2.8 |
| S3: AC150, $\beta$ dose 25 Gy | Small | 0.7 | (0.0–1.8) | 3.3 | 3.4 | 1.6 | (0.7–3.0) | 3.4 | 3.8 |
| | Large | 0.7 | (0.3–1.2) | 1.2 | 1.4 | 1.6 | (1.1–2.4) | 1.3 | 2.1 |
| S4: AC150, field dose $\sim$ 25 Gy | Small | 0.9 | (0.0–2.0) | 2.4 | 2.6 | 7.1 | (5.2–10.0) | 2.4 | 7.5 |
| | Large | 0.9 | (0.5–1.6) | 1.4 | 1.6 | 3.7 | (2.8–5.2) | 1.3 | 3.9 |
| S5: K162, $\beta$ dose, 2.74 Gy | Small | 0.8 | (0.0–2.9) | 5.5 | 5.6 | 11.5 | (8.2–16.6) | 5.1 | 12.6 |
| | Large | 0.8 | (0.0–1.5) | 1.4 | 1.6 | 3.1 | (2.3–4.4) | 1.4 | 3.4 |
| S6: K162, field dose $\sim$ 46 Gy | Small | 2.6 | (1.7–4.2) | 3.7 | 4.5 | 17.7 | (13.6–24.3) | 3.3 | 18.0 |
| | Large | 2.5 | (1.8–3.5) | 1.9 | 3.1 | 11.8 | (9.0–16.3) | 1.7 | 11.9 |
| Combined estimate of $\sigma_b$ for $R_1$ vs. $R_2$ | | 1.4 | (1.2–1.6) | | | | | | |

# Estimates of over-dispersion

- Up to 3% for recycling ratios
- Several % for dose recovery tests (surrogate naturals)
- More for natural samples (typically 10–20%)

# Age models (actually, $D_e$ models)

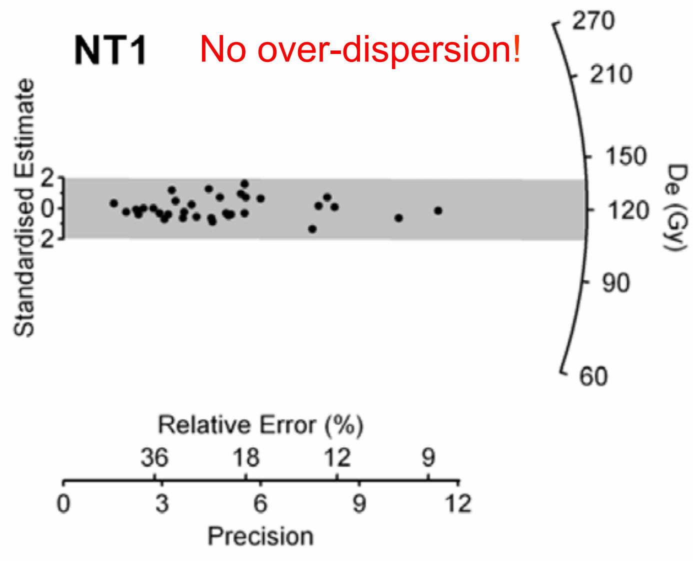
- Typically, the weighted mean of $Y_1$, $Y_2$ … etc. is calculated as:

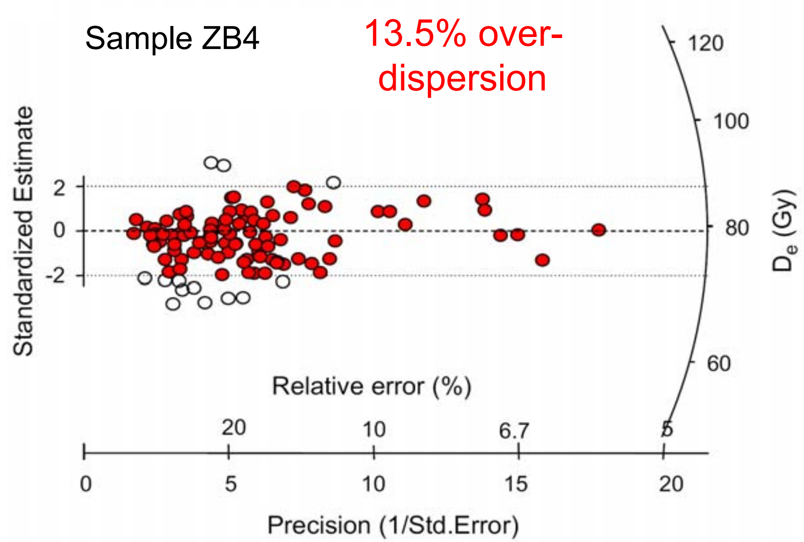
$$\bar{Y} = \frac{Y_1/\sigma_1^2 + Y_2/\sigma_2^2 + \ldots}{1/\sigma_1^2 + 1/\sigma_2^2 + \ldots}$$

- Note that the weights are by absolute (not relative) standard error
- So the weighted mean of $D_e$ values with the same relative errors will be biased towards the smaller values
- Now, the relative standard error of an estimate ≈ absolute standard error of its natural logarithm ($\log_e$)
- So, by using the natural logs of the $D_e$ values, the weighted mean can be calculated with the relative standard errors
- The ‘Central’ and ‘Common’ age models do this

# Central and Common age models

- Developed for fission-track ages by Galbraith & Laslett (*NTRM* 1993)
- Extended to $D_e$ estimates by Galbraith *et al*. (*Archaeometry* 1999)
- Both give weighted mean (≈ geometric mean of true $D_e$ values)
- Central age model calculates and includes any over-dispersion



# When (not) to use these models

- Central/Common age models <u>may</u> be appropriate for sediments that:
  - were well bleached when deposited
  - have remained undisturbed since burial
  - and have no dosimetry complications (‘hot’ spots, ‘cold’ spots)

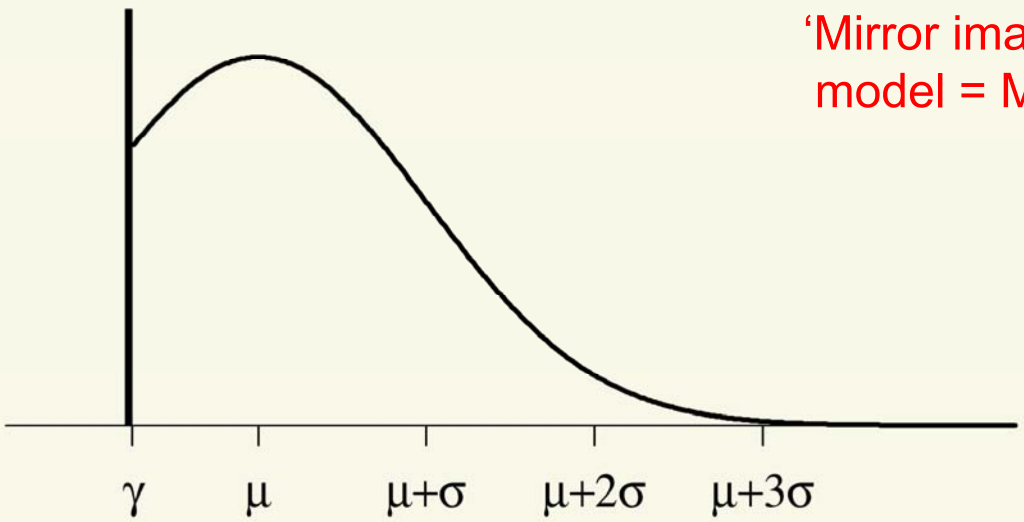
- Probably <u>not</u> suitable for poorly bleached and/or mixed samples

- Alternative models include:
  - Minimum age model (Galbraith & Laslett, Galbraith *et al*.)
  - Maximum age model (Olley *et al*., *Quat Sci Rev* 2006)
  - Finite mixture model (Galbraith & Green, *NTRM* 1990; Roberts *et al*., *Radiat Meas* 2000)
- Key points:
  - field context is vital
  - take <u>all</u> available evidence into account to decide model choice
  - obtain independent estimate of over-dispersion for MAM & FMM
  - don’t apply every model to every sample!

# Continuous mixtures

- Where the smallest $D_e$ values ≈ age of 'target' event
- 3- and 4-parameter 'Minimum' age models:

Here $x_i$ has a mixture distribution — a mixture of a single minimum value $\gamma$ with probability $p$ and a continuous range of values greater than $\gamma$, usually modelled as a truncated normal distribution.

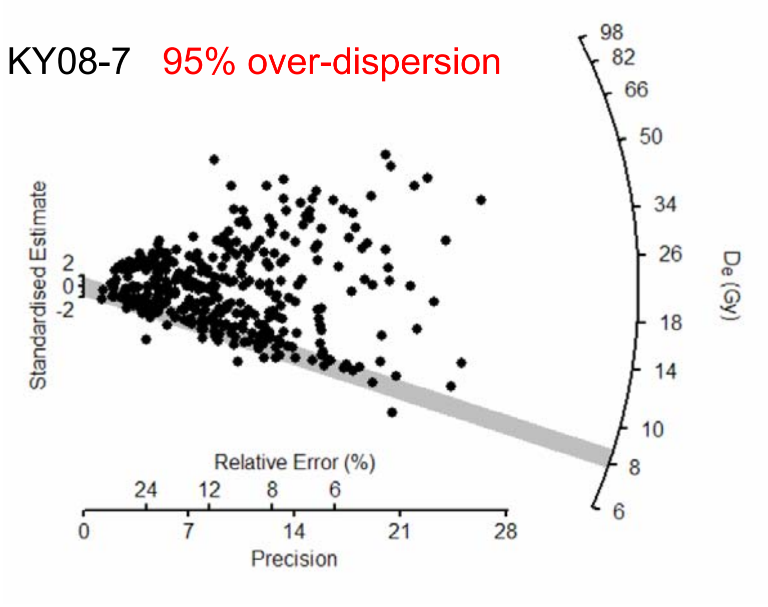



'Mirror image' of Minimum age model = Maximum age model

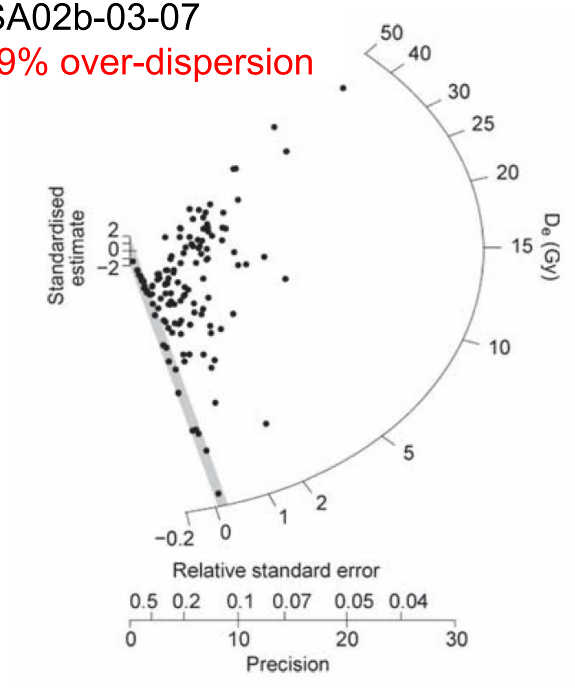
There are 4 parameters: $\gamma, \mu, \sigma$ and $p$, the proportion of grains that have the minimum dose. $\gamma = \mu$ in 3-parameter version.

# Minimum age model

- Add an estimate of over-dispersion to each measurement error (in quadrature) <u>before</u> running model
- Standard (logged) version of model appropriate for most natural samples
- Modern or young samples, use <u>un-logged</u> version (Arnold *et al*., *Quat Geochron* 2009)
- Check fits using formal statistical criteria (e.g. maximum log likelihood)

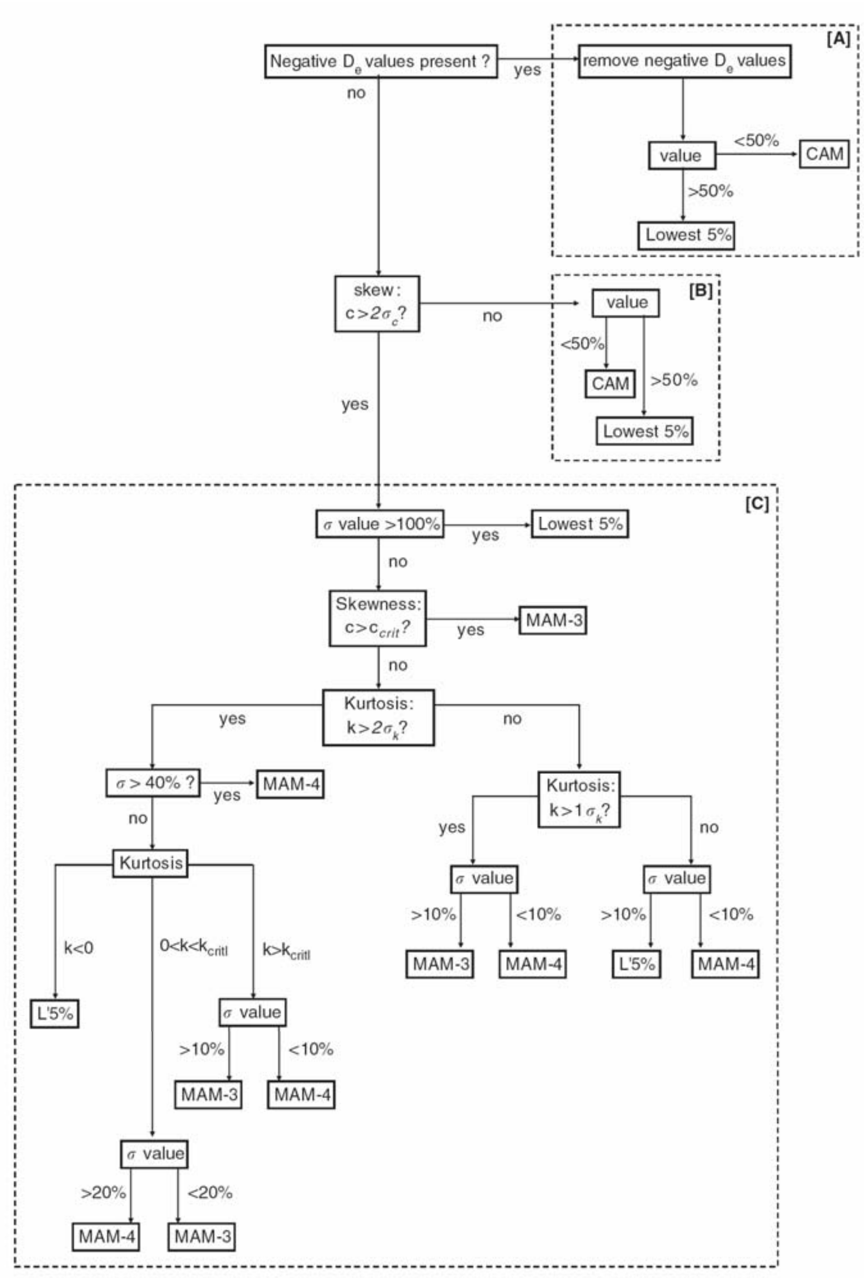



USA02b-03-07
119% over-dispersion



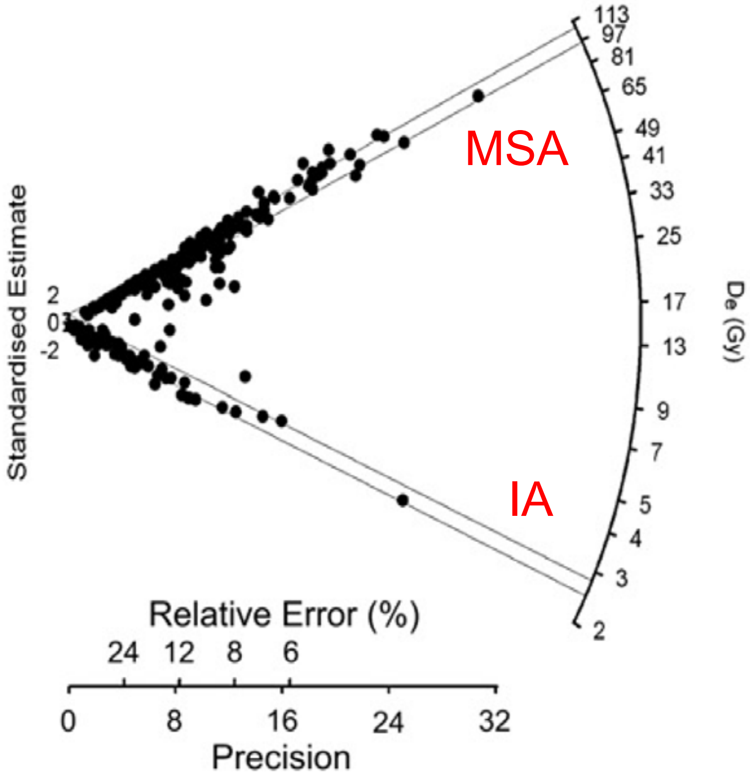
# Flowcharts – friend or foe?

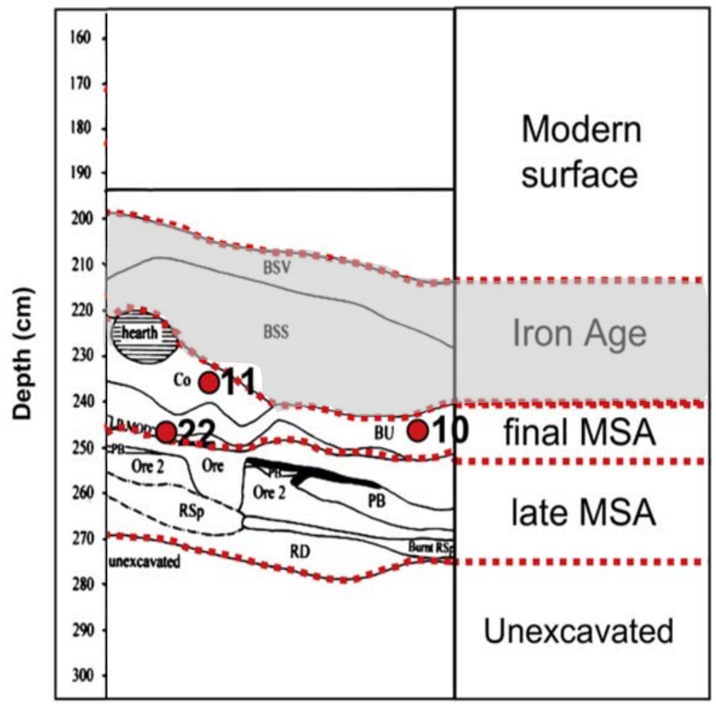
- Model decision based on simulated data (Bailey & Arnold, *Quat Sci Rev* 2006)
- Validity depends on what data went into model
- 'Real world' complications, e.g. between-sample variation in over-dispersion
- Skewness and kurtosis:
  - conflate 2 sources of error
  - should be logged values
- Don't rely on flowcharts



## Discontinuous mixtures

- ‘Finite’ mixture model for samples composed of discrete age-populations of grains mixed together after burial
  - common at archaeological sites (Jacobs & Roberts, *Evol Anthrop* 2007)
  - e.g. sample 11 from Sibudu Cave (South Africa) is a mixture of Iron Age (IA) and Middle Stone Age (MSA) sediments



# Finite mixture model

- Apply <u>only</u> to single grains and <u>not</u> to continuous mixtures!

- **Input:**
  - single-grain $D_e$ values and standard errors
  - estimated number of finite $D_e$ components
  - estimate of inherent over-dispersion ($\sigma_b$)

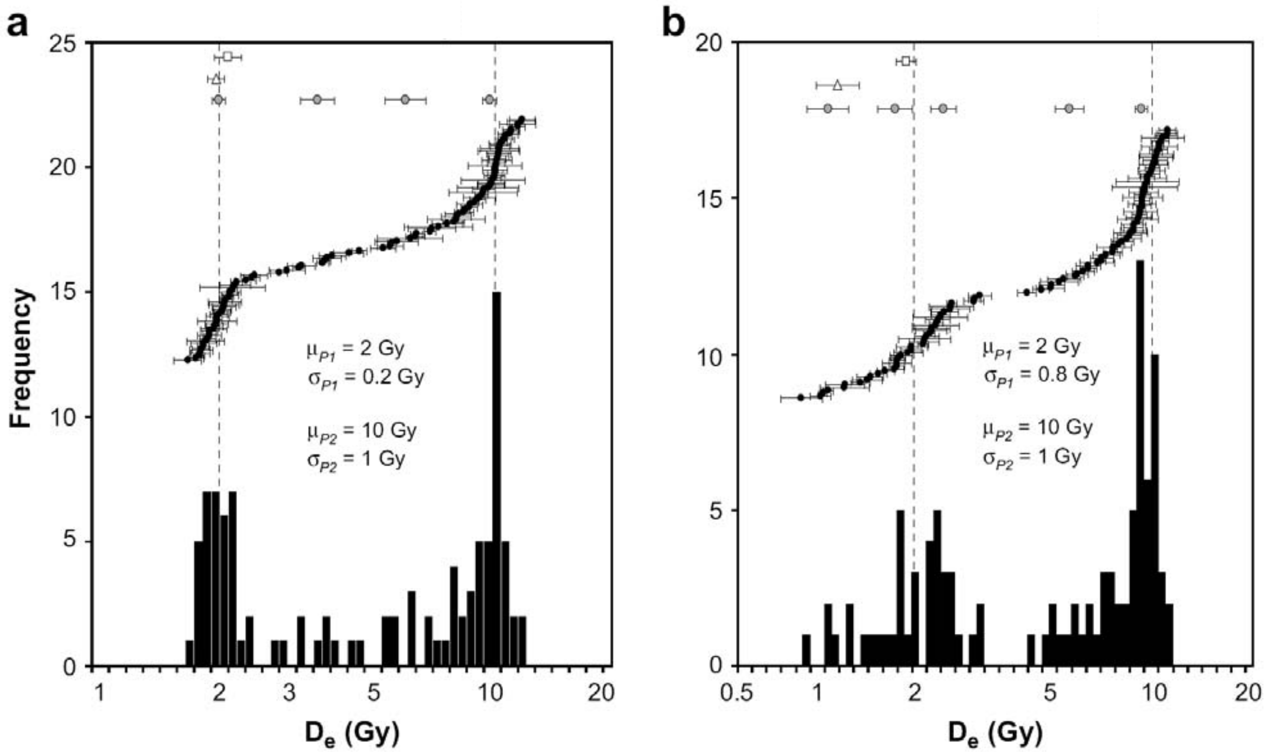
- **Output:**
  - $D_e$ and standard error of each component (Central age model)
  - relative proportion of grains in each component
  - two estimates of goodness-of-fit
    - maximum log likelihood (MLL)
    - Bayes Information Criterion (BIC)

- FMM developed for fission-track ages by Galbraith & Green (1990)

- But fission-track ages have zero over-dispersion, so OSL version (Roberts *et al*., 2000) includes over-dispersion as <u>extra</u> parameter

- **Procedure:**
  - change number of components and over-dispersions to:
    - maximise MLL (increase by > 2 for each added component)
    - minimise BIC (MLL penalised for each added component)
  - MLL and BIC may not always give same best-fit outcomes, so check result not sensitive to different over-dispersion values
    - David *et al*. (*J Quat Sci* 2007), Jacobs *et al*. (*J Arch Sci* 2008)

FMM will detect 'phantom' $D_e$ populations even for very small aliquots, e.g. 4-grain aliquots with only 1–2 luminescent grains (Arnold & Roberts, *Quat Geochron* 2009)



# Concluding remarks

- Models are garbage collectors, so collect good data!
- Correctly estimate your measurement errors
- Display your $D_e$ values meaningfully (as radial plots)
- Measure (and report) your $D_e$ over-dispersion values
- Consider sample context when selecting an age model
- Include over-dispersion when implementing age model
- Don’t use FMM for m-g aliquots or continuous mixtures