ON THE VISUALISATION AND INTERCOMPARISON OF DETRITAL AGE DISTRIBUTIONS

PART 1: VISUALISATION

Pieter Vermeesch

London Geochronology Centre
University College London
p.vermeesch@ucl.ac.uk

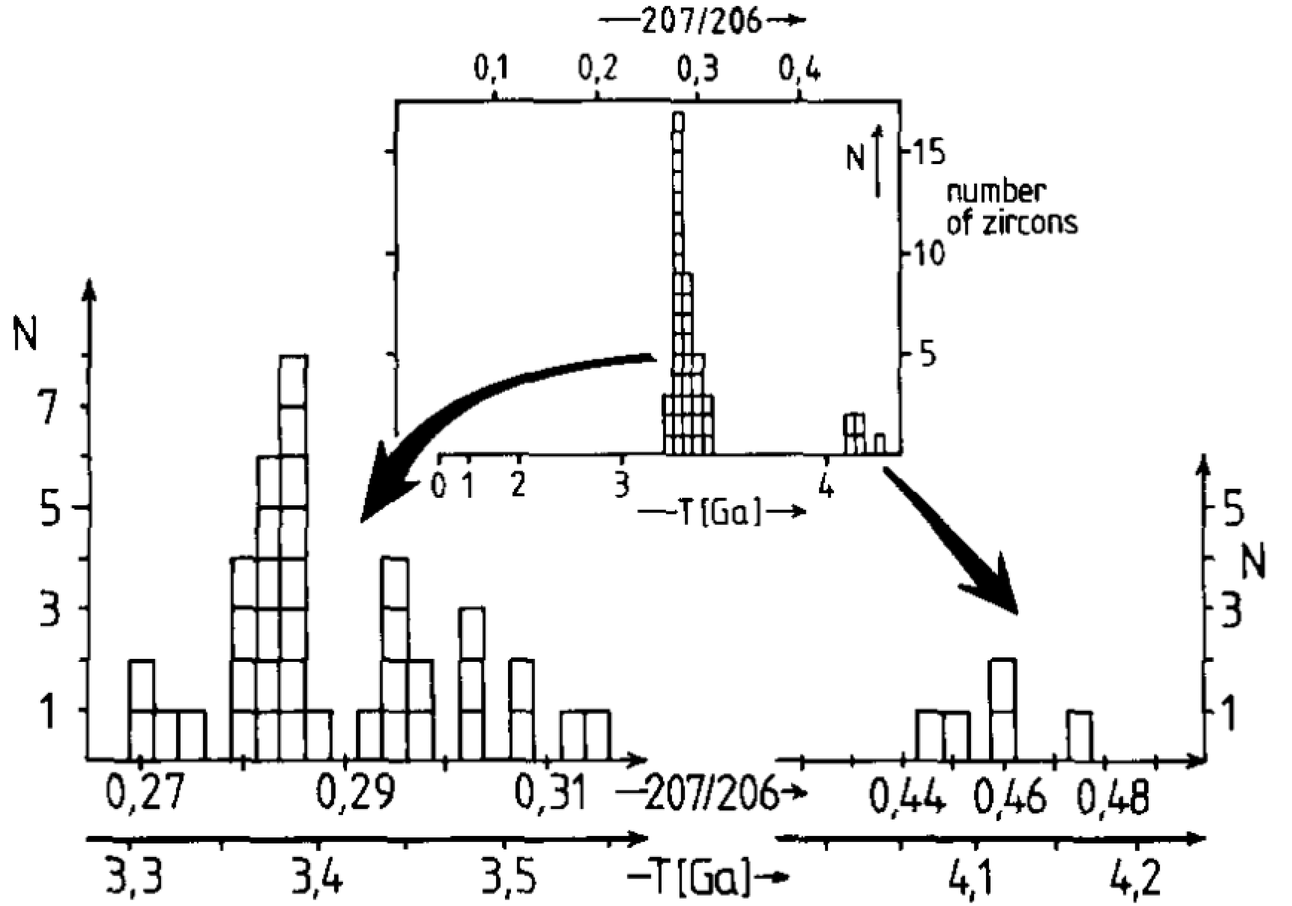
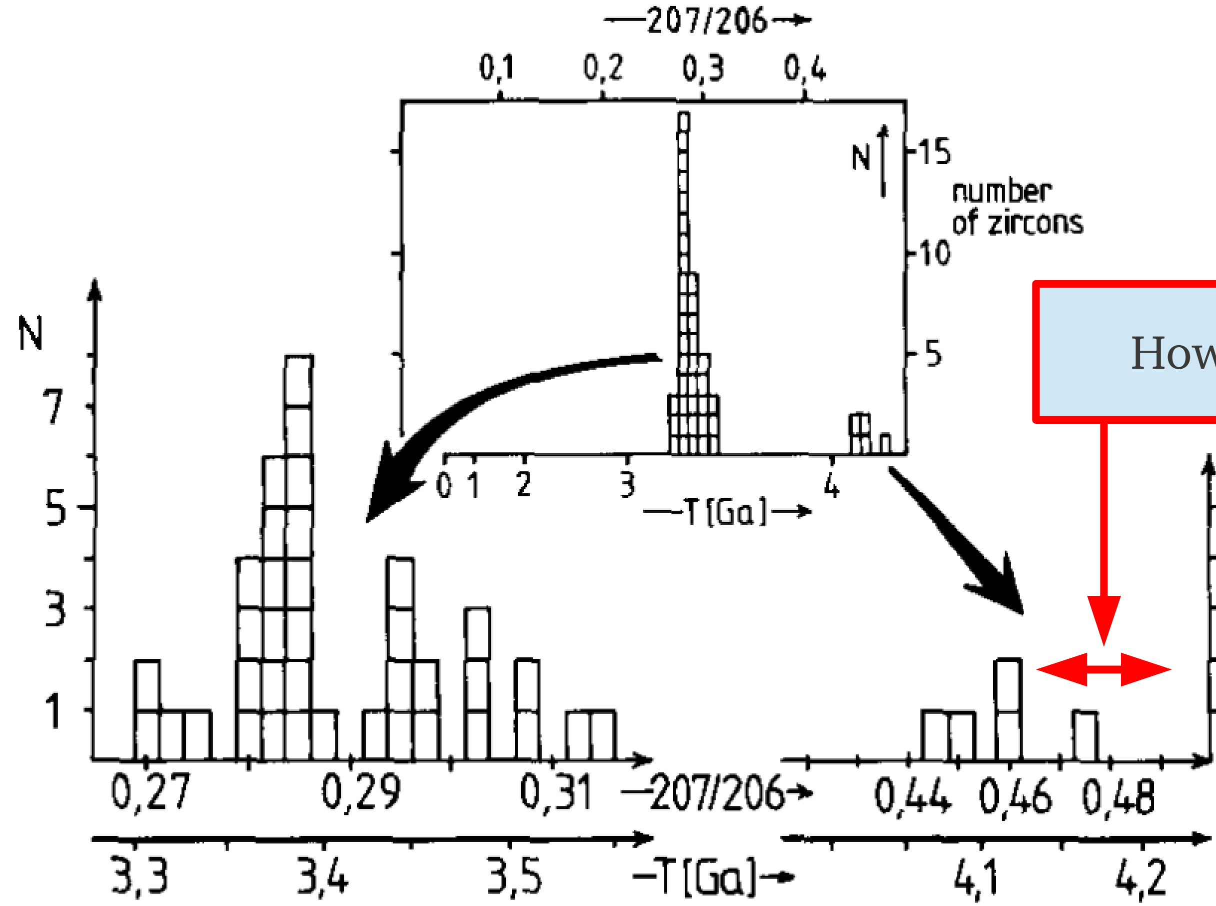


Fig. 4. Frequency histogram of 207/206 ages determined for 42 zircons by the evaporation method. The variations in the age distributions of both the Archean zircon subpopulations are far outside the routine data reproducibility of 1%. Widths of the apparent ages classes: $\Delta(207/206) = 0.01$ (inset), $\Delta(207/206) = 0.0025$ (expanded presentation).



How to place the bins?

Fig. 4. Frequency histogram of 207/206 ages determined for 42 zircons by the evaporation method. The variations in the age distributions of both the Archean zircon subpopulations are far outside the routine data reproducibility of 1%. Widths of the apparent ages classes: $\Delta(207/206) = 0.01$ (inset), $\Delta(207/206) = 0.0025$ (expanded presentation).

$k = \sqrt{n}$ (Excel)

$k = [\log_2 n + 1]$ (Sturges' Rule)

$h = 2\frac{IQR(x)}{n^{1/3}}$ (Freedman–Diaconis)

...

A search for ancient detrital zircons in Zimbabwean sediments

M. H. DODSON,[1*] W. COMPSTON,[1] I. S. WILLIAMS[1] & J. F. WILSON[2]

Journal of the Geological Society 1988, v.145; p977-983.

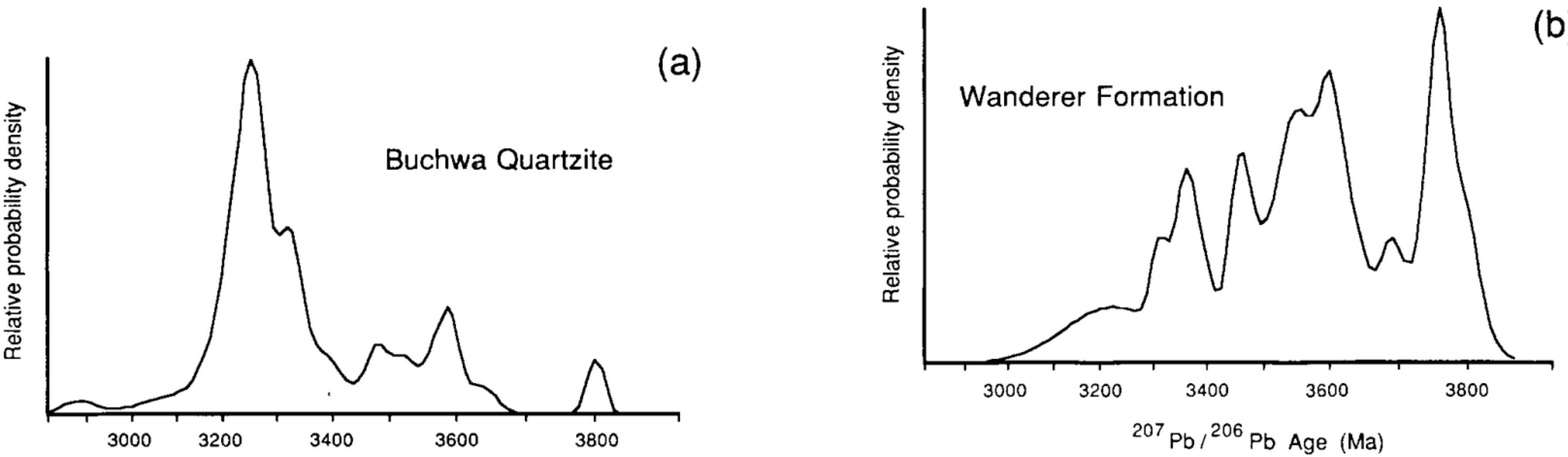
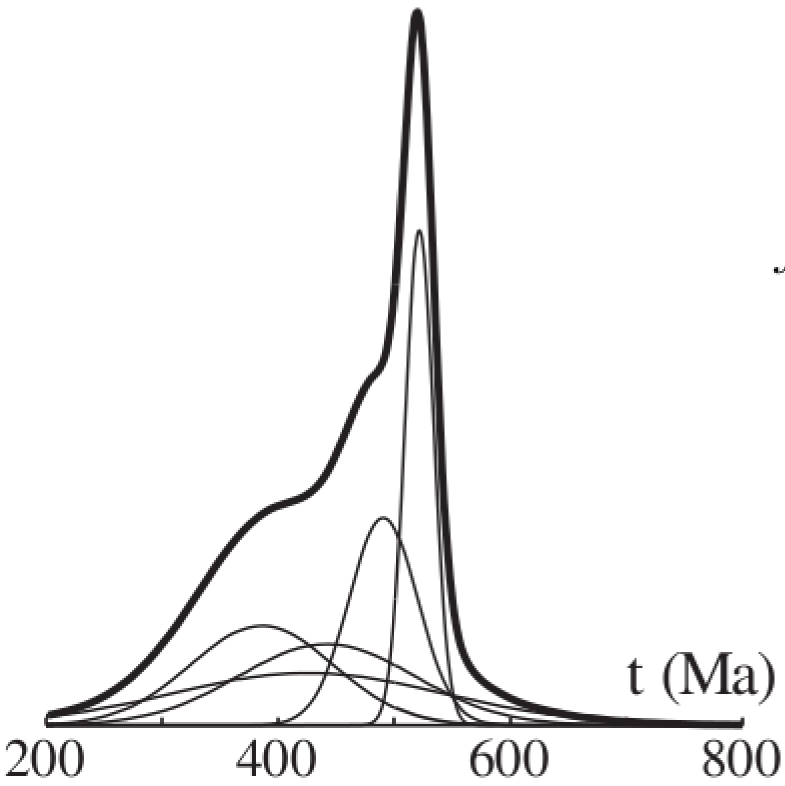


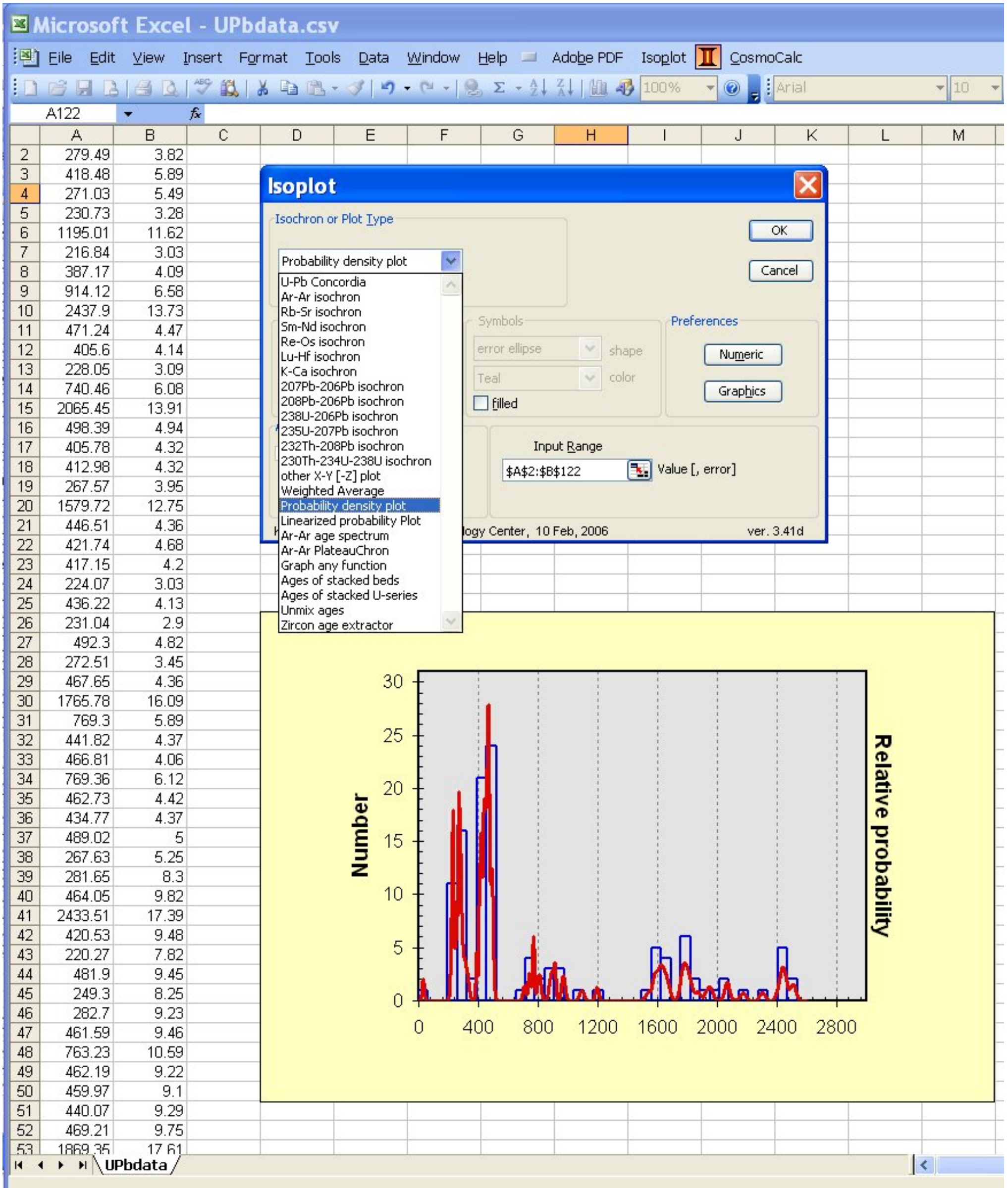
Fig. 3. 'Histograms' of $^{207}Pb/^{206}Pb$ ages of zircons extracted from **(a)** Buchwa Quartzite and **(b)** Wanderer Formation. The curves are obtained by summing Gaussian distributions of constant area, one for each age with standard deviation equal to the calculated uncertainty in the age. (Technique and program due to Dr P. Zeitler, ANU.)

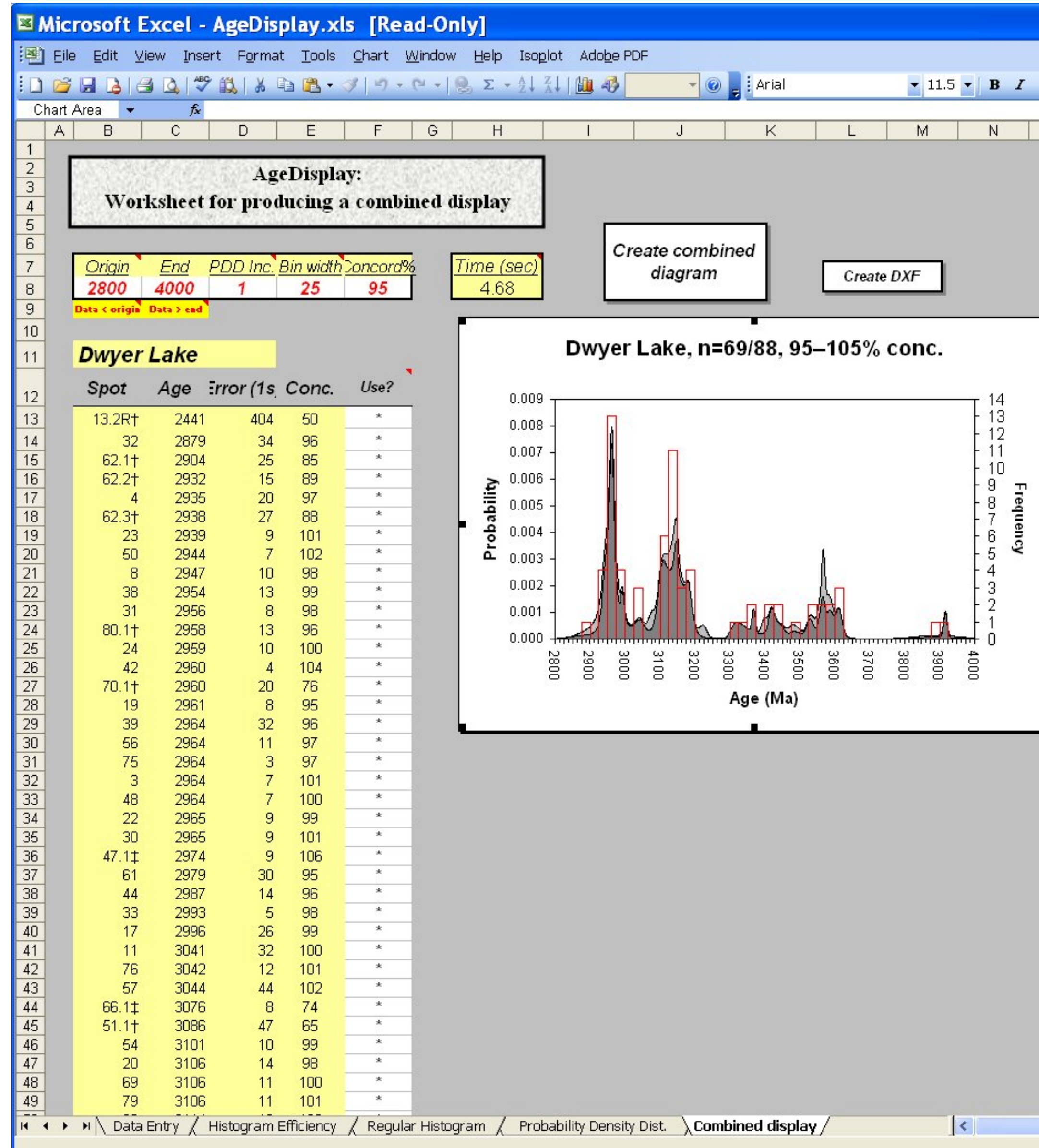


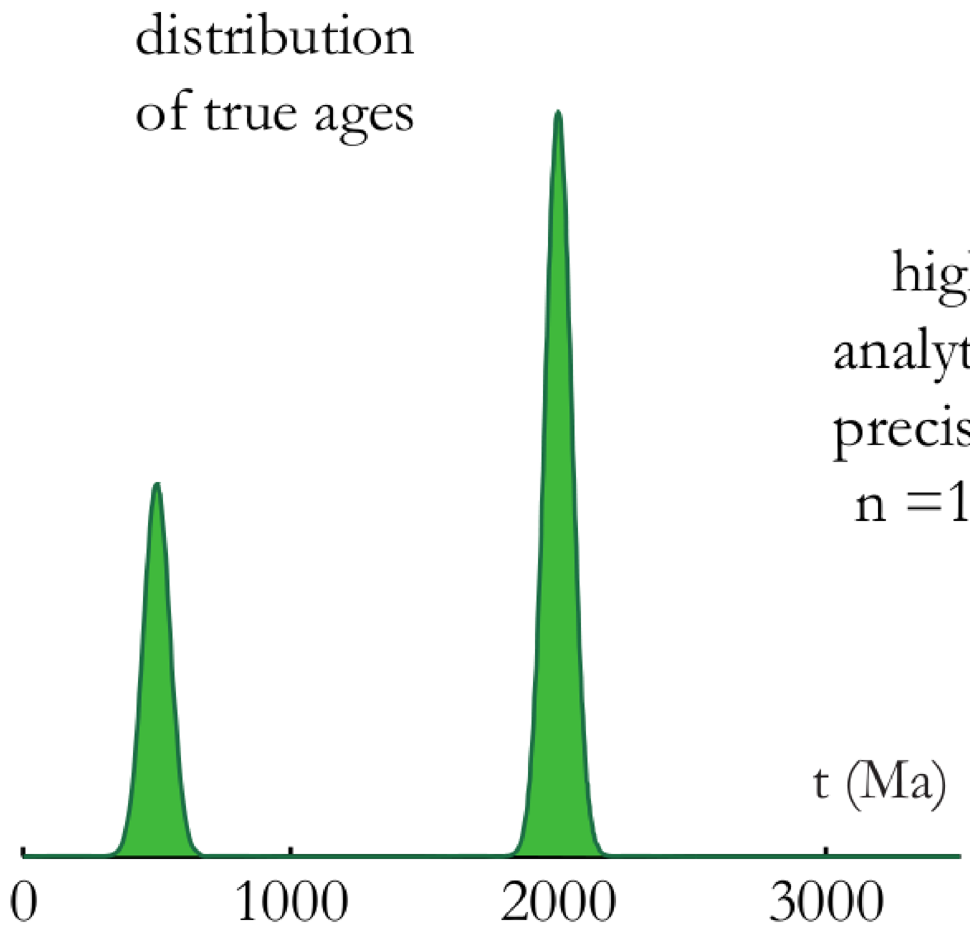
$$f(t) = \sum_{i=1}^{N} \frac{1}{e_i\sqrt{2\pi}} \exp^{-(t-x_i)^2/2e_i^2}$$

Isoplot (Ludwig, 2003)

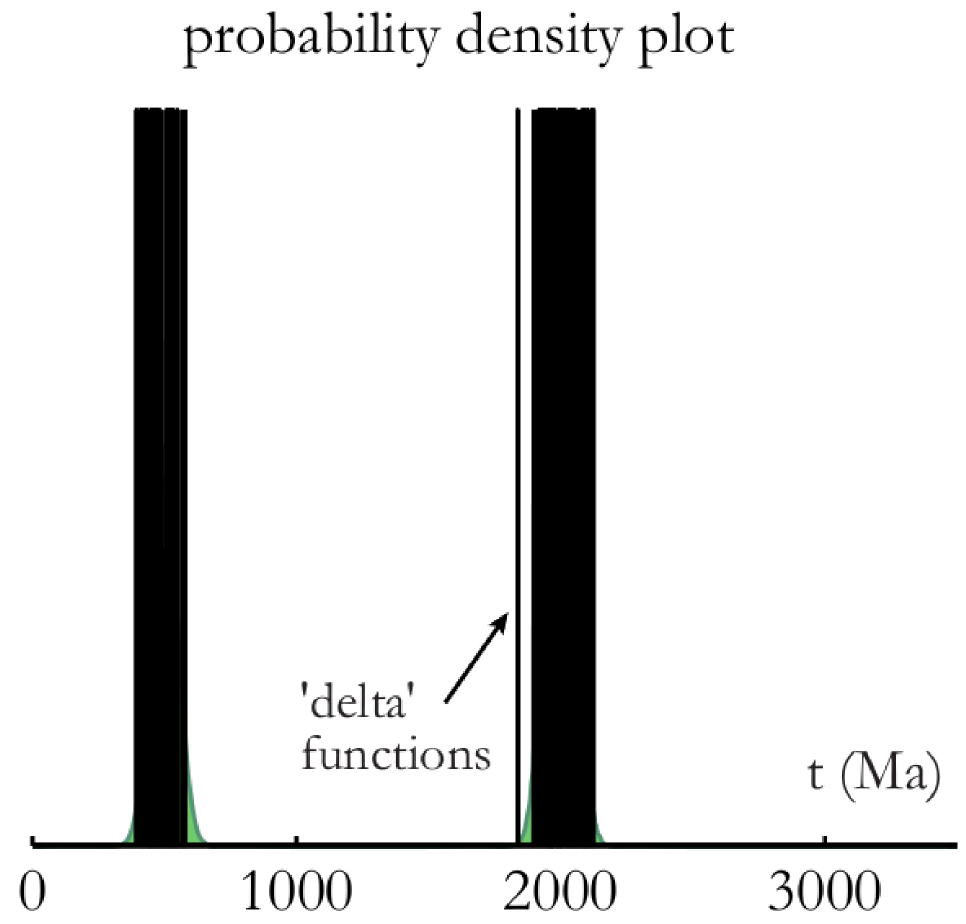
AgeDisplay (Sircombe, 2004)

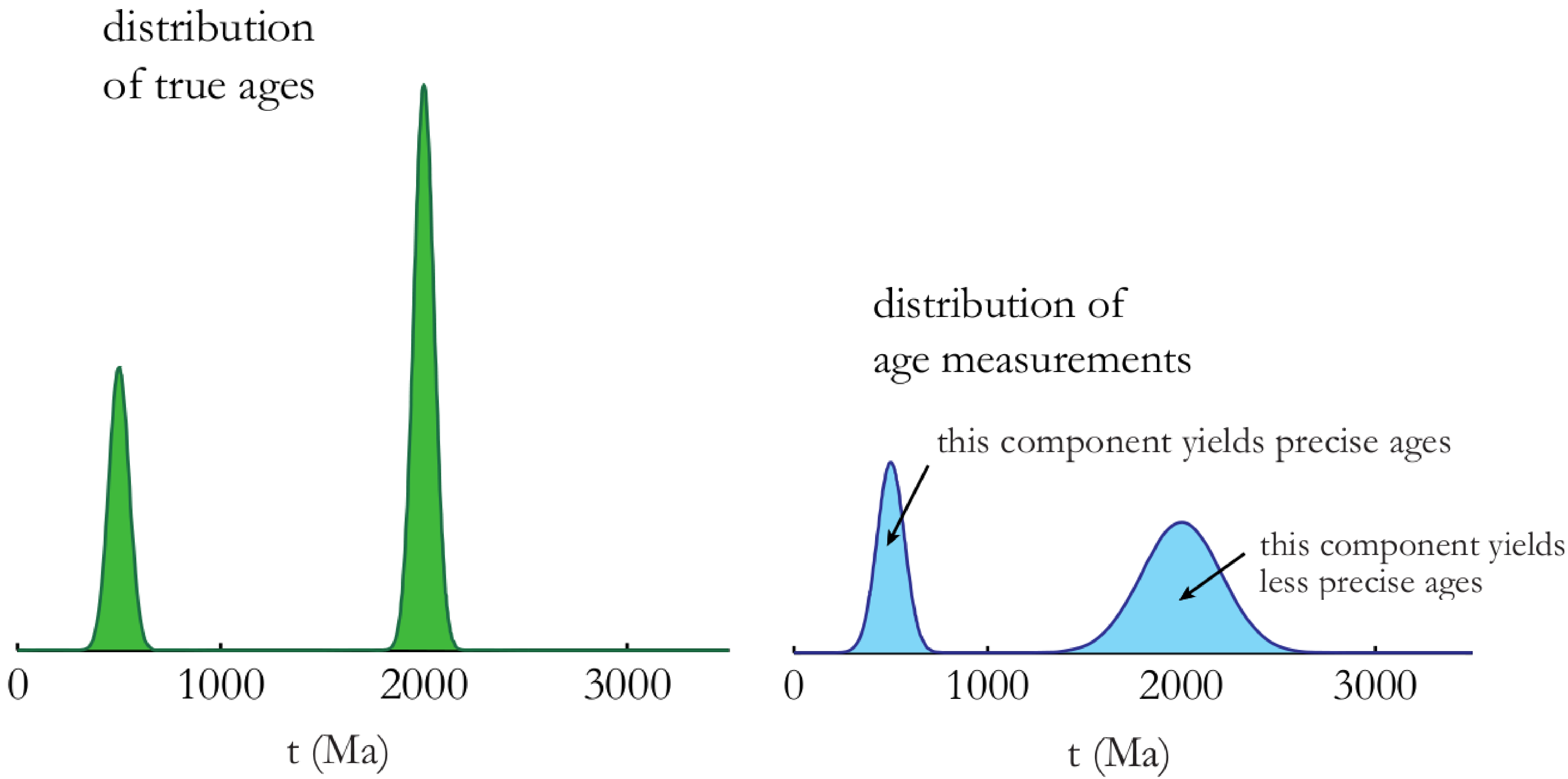


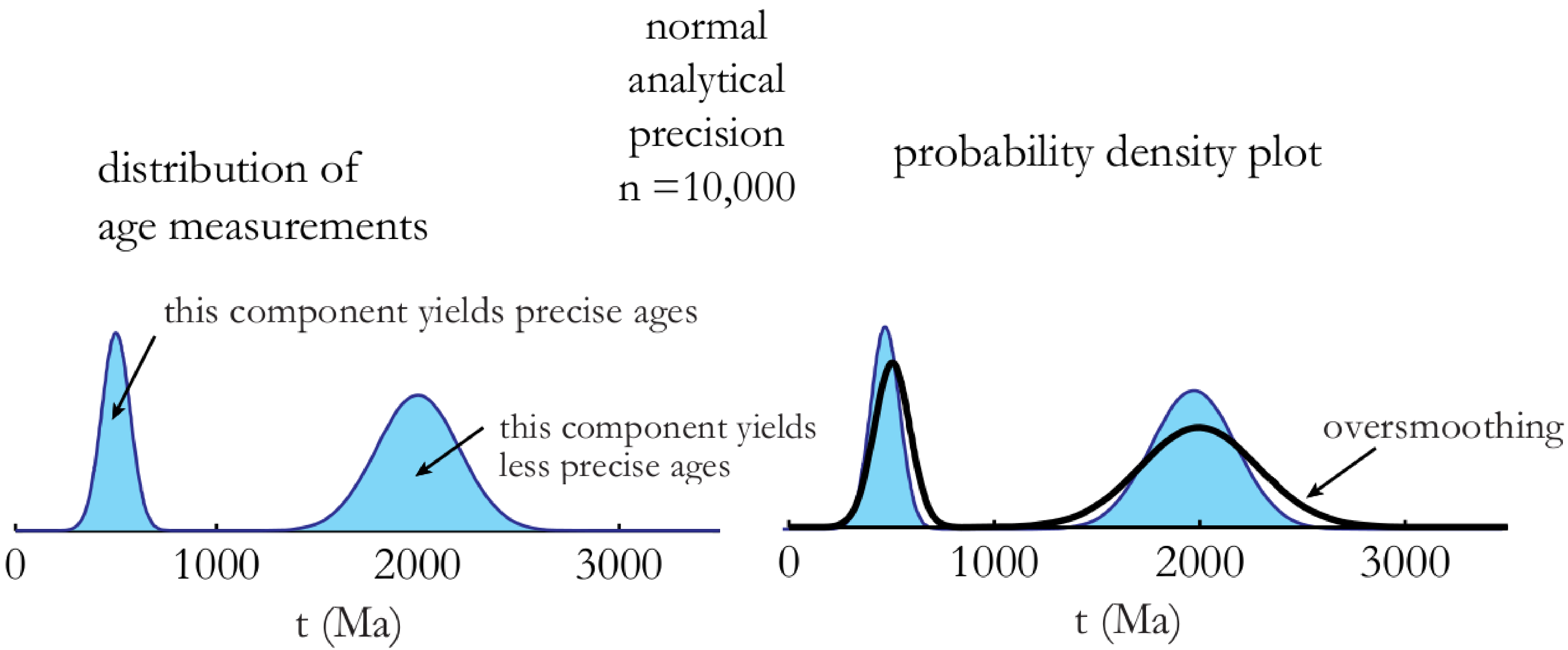




high
analytical
precision
n =117



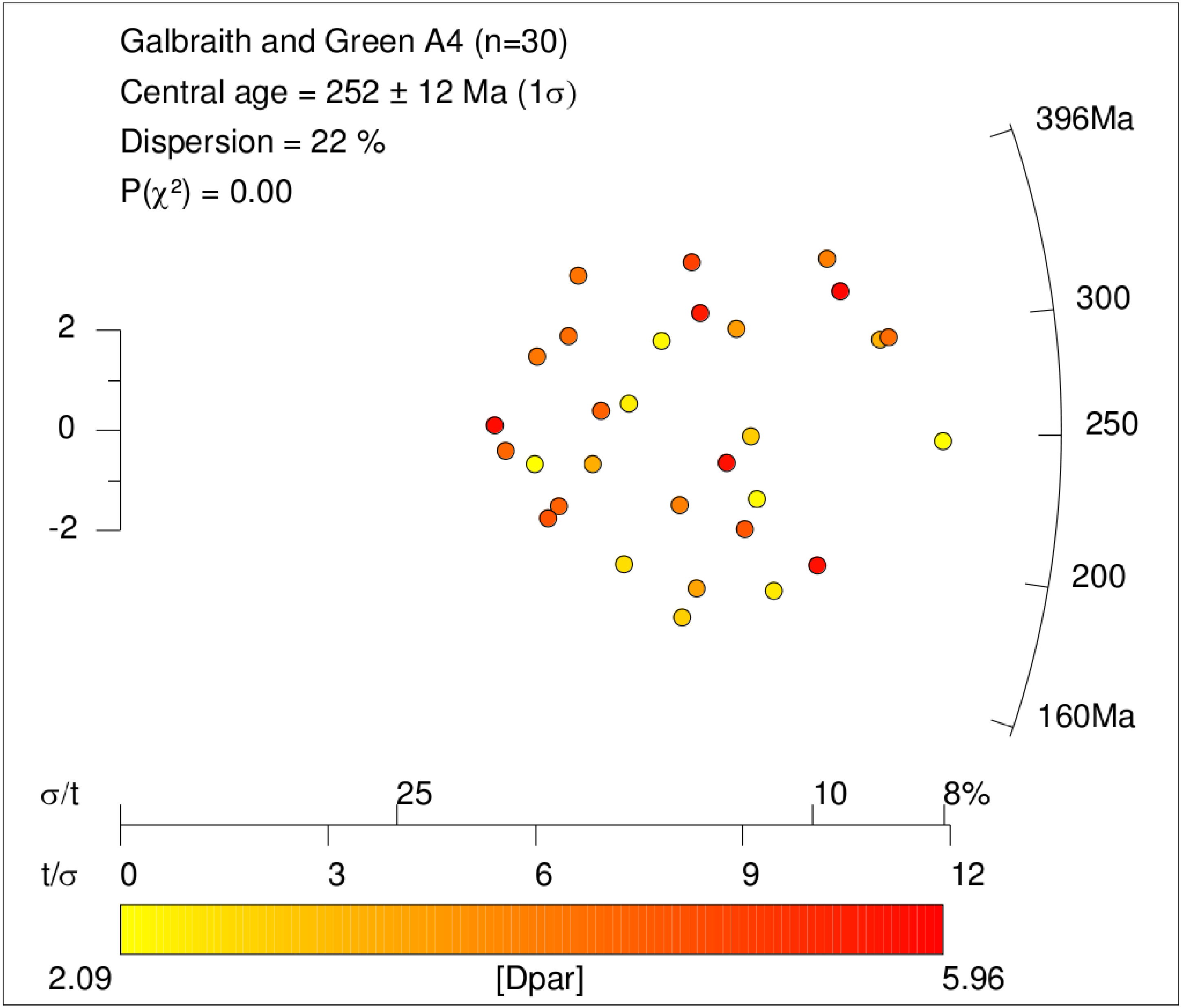

distribution
of true ages
0
1000
2000
3000
t (Ma)
distribution of
age measurements
this component yields precise ages
this component yields
less precise ages
0
1000
2000
3000
t (Ma)


normal
analytical
precision
n =10,000
distribution of
age measurements
this component yields precise ages
this component yields
less precise ages
0
1000
2000
3000
t (Ma)
probability density plot
oversmoothing
0
1000
2000
3000
t (Ma)

DEFINITION: HOMOSCEDASTICITY
If the variance about the means of two samples is equal, this condition is called *homoscedasticity*. If the variances are unequal, the two samples are termed *heteroscedastic*.

Burt, James E.
Elementary statistics for geographers / James E. Burt, Gerald M. Barber, David L. Rigby. — 3rd ed.

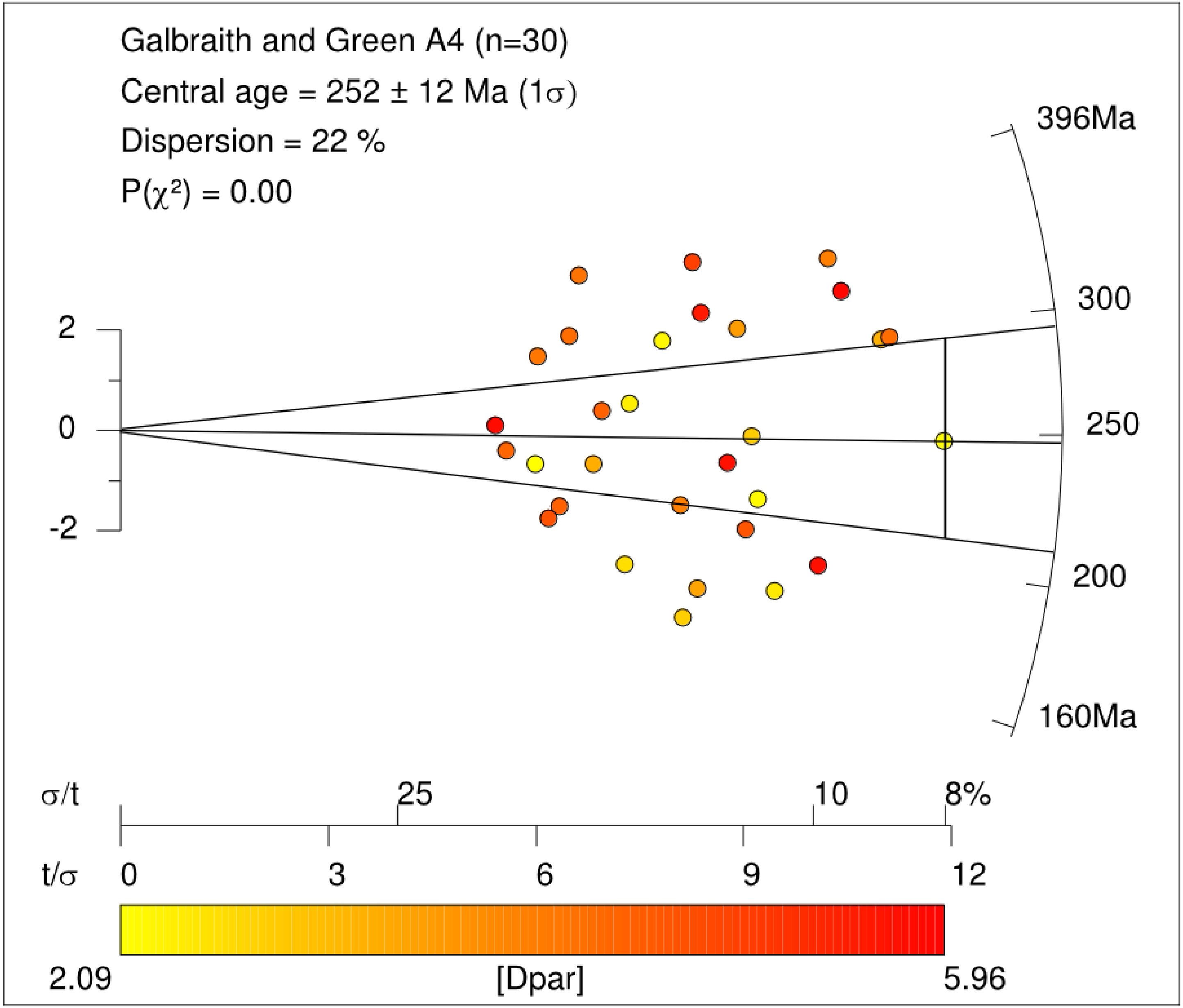
Radial plots are the only way to simultaneously visualise age and uncertainty

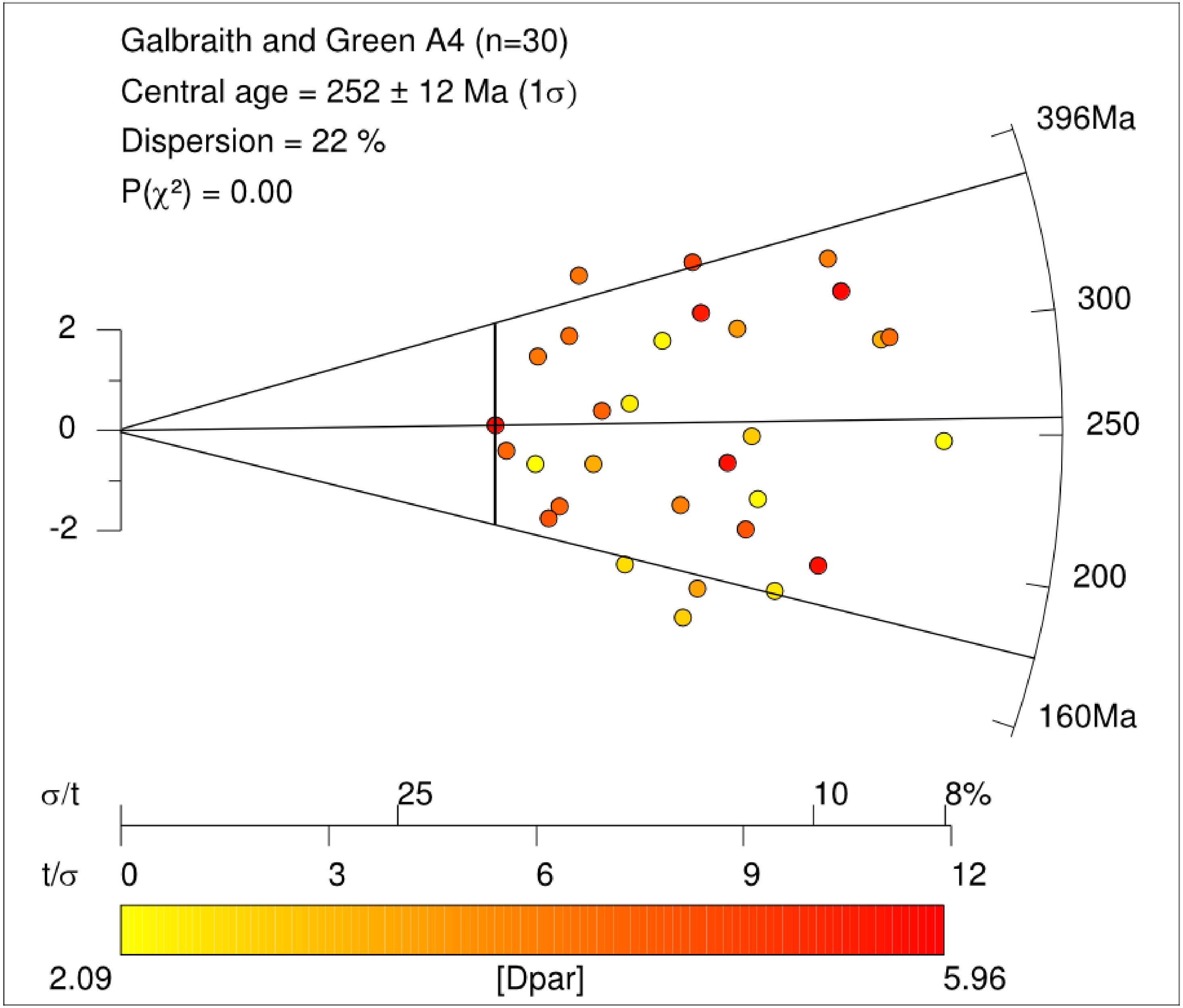


$x_j = 1/\sigma(z_j)$

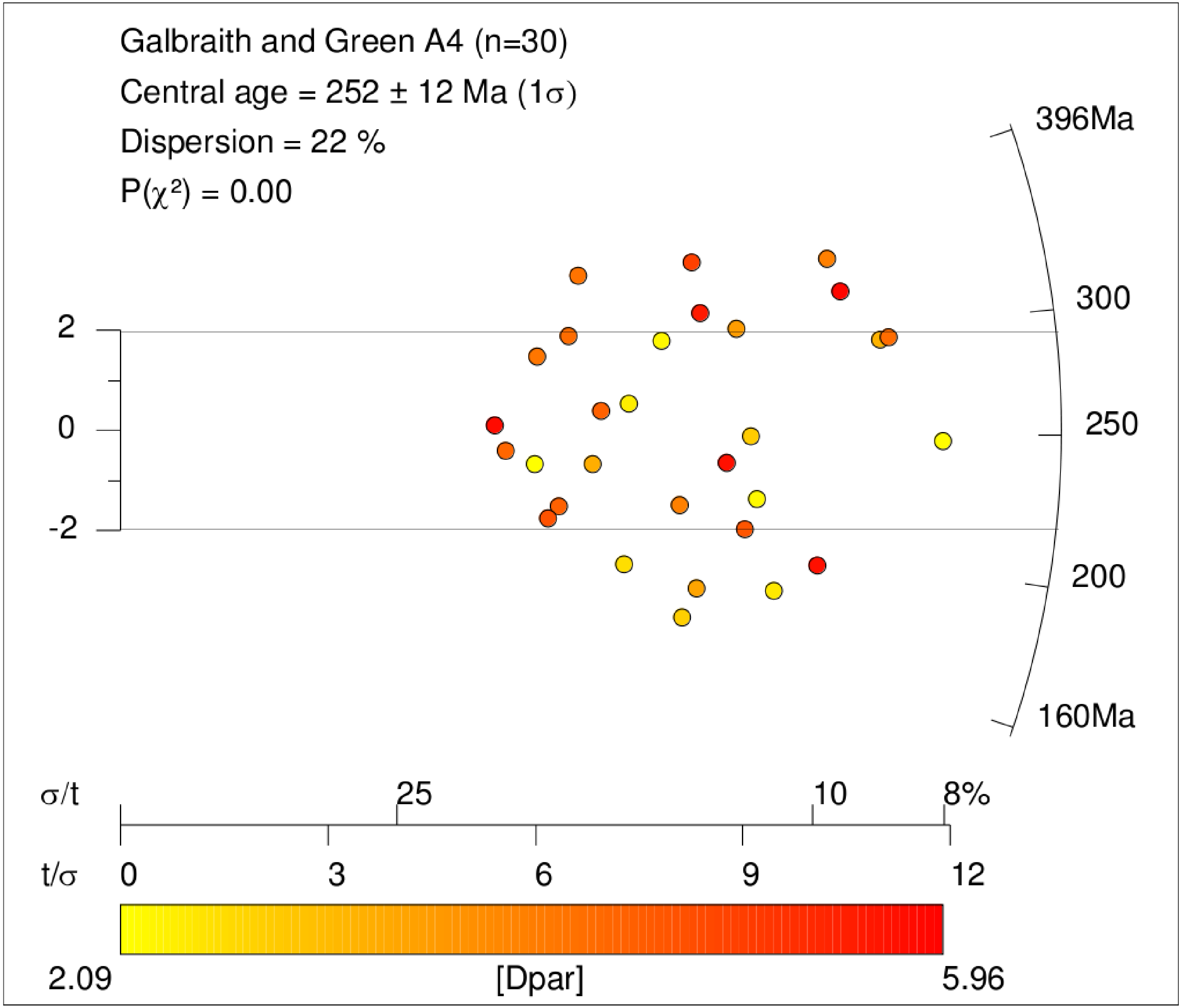
$y_j = (z_j - z_o)/\sigma(z_j)$

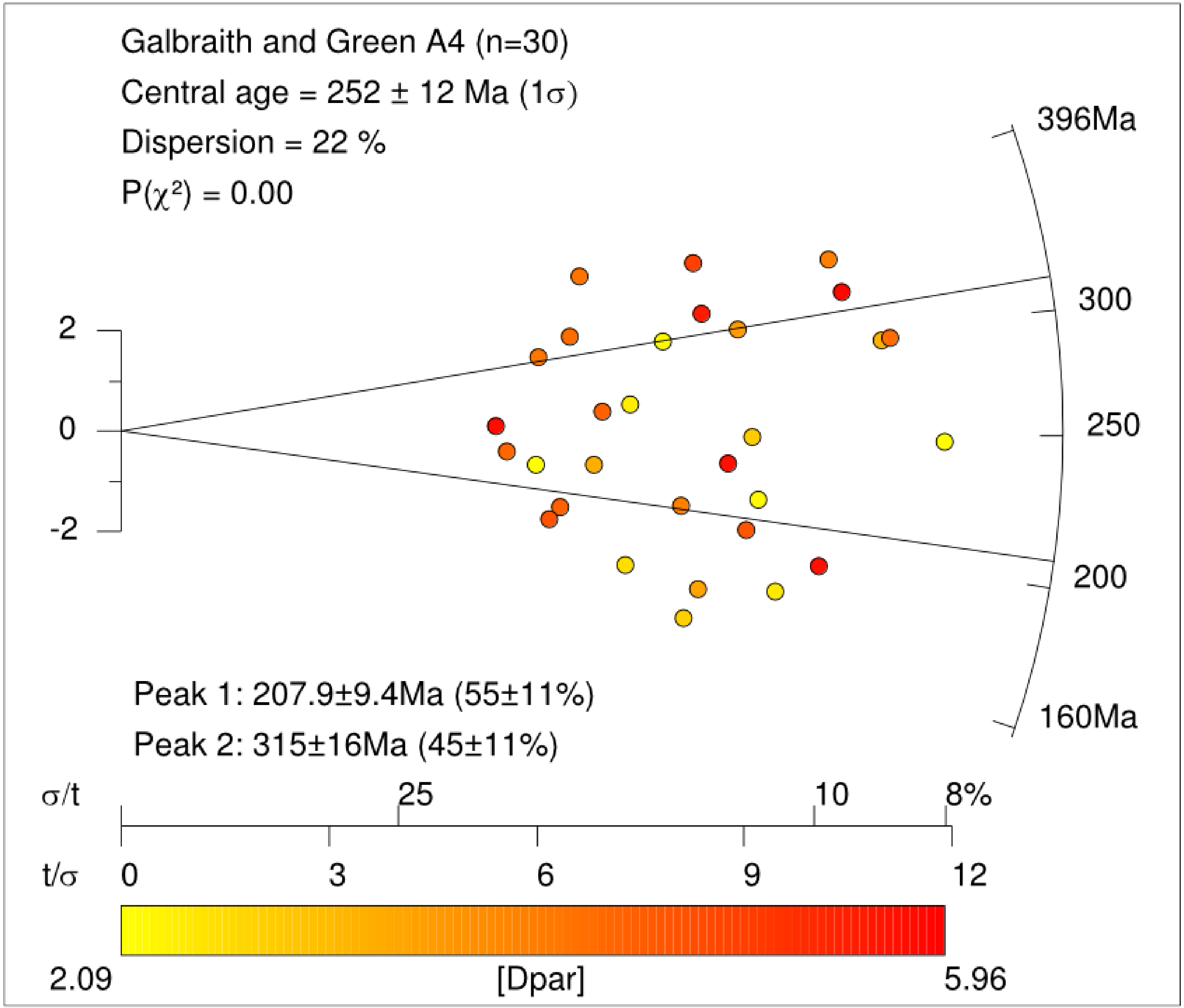
Galbraith
(*Technometrics*, 1988)


Galbraith and Green A4 (n=30)
Central age = 252 ± 12 Ma (1σ)
Dispersion = 22 %
P(χ²) = 0.00
396Ma
300
250
200
160Ma
2
0
-2
σ/t
25
10
8%
t/σ
0
3
6
9
12
2.09
[Dpar]
5.96

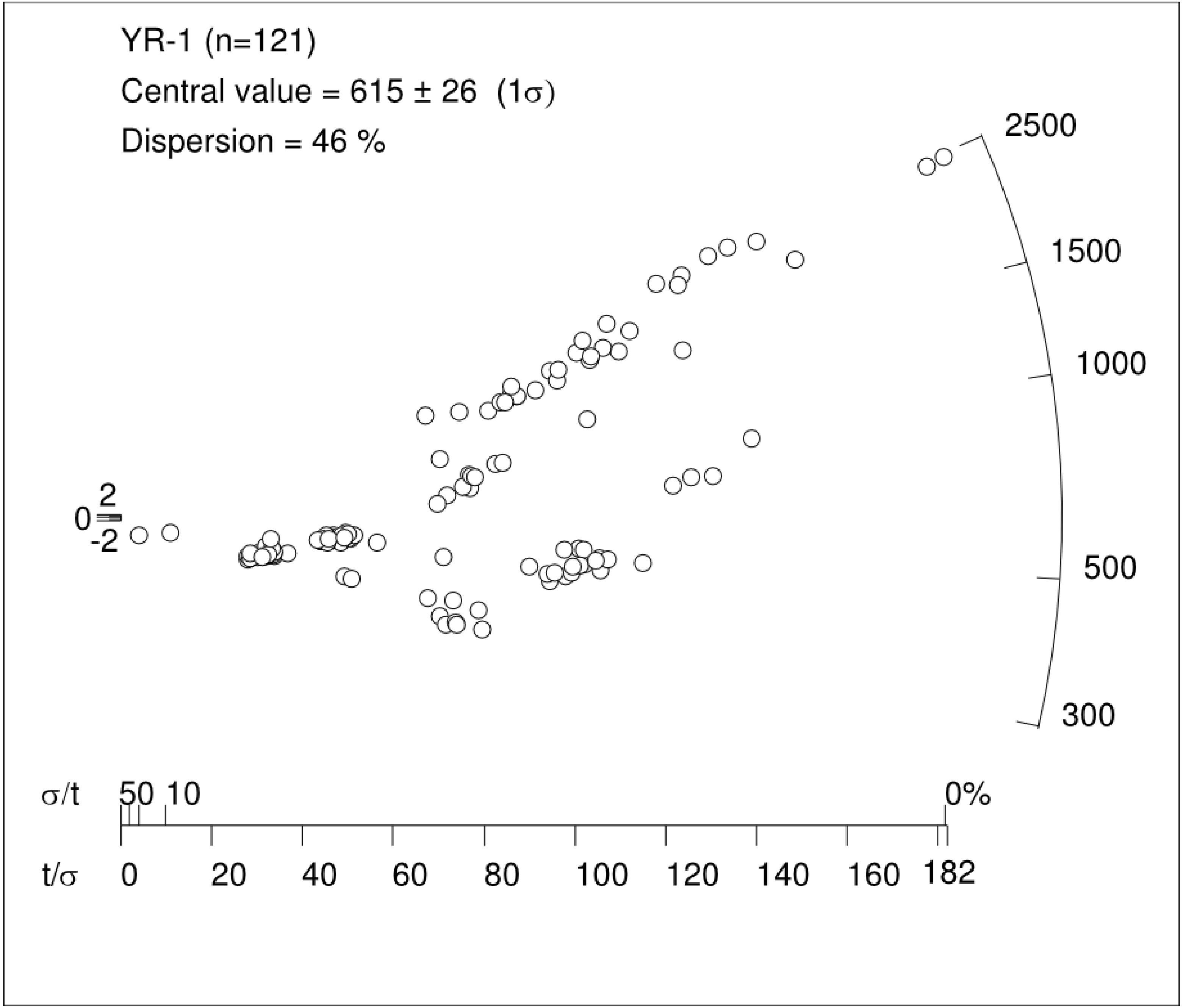


Galbraith and Green A4 (n=30)
Central age = 252 ± 12 Ma (1σ)
Dispersion = 22 %
P(χ²) = 0.00
396Ma
300
250
200
160Ma
2
0
-2
σ/t
25
10
8%
t/σ
0
3
6
9
12
2.09
[Dpar]
5.96

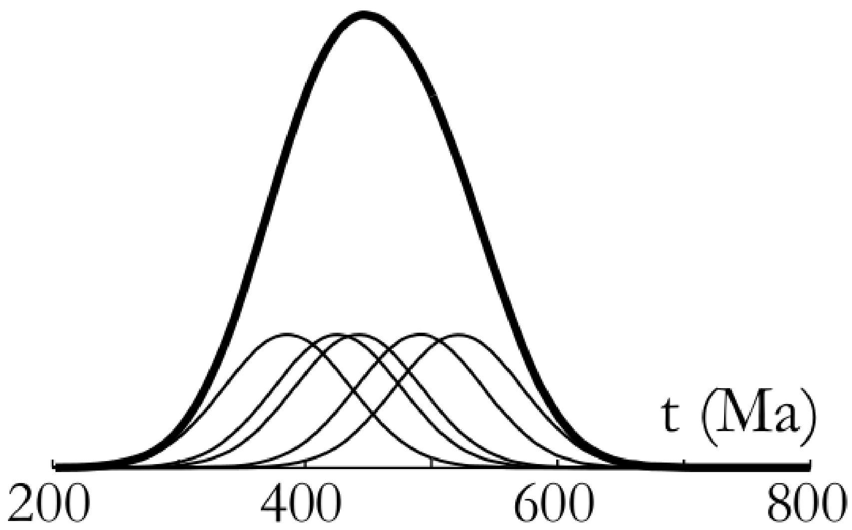

Galbraith and Green A4 (n=30)
Central age = 252 ± 12 Ma (1σ)
Dispersion = 22 %
P(χ²) = 0.00
396Ma
300
250
200
160Ma
2
0
-2
σ/t
25
10
8%
t/σ
0
3
6
9
12
2.09
[Dpar]
5.96


Galbraith and Green A4 (n=30)
Central age = 252 ± 12 Ma (1σ)
Dispersion = 22 %
P(χ²) = 0.00
396Ma
300
250
200
160Ma
2
0
-2
Peak 1: 207.9±9.4Ma (55±11%)
Peak 2: 315±16Ma (45±11%)
σ/t
25
10
8%
t/σ
0
3
6
9
12
2.09
[Dpar]
5.96

Radial plots may not be very useful for detrital U-Pb age distributions

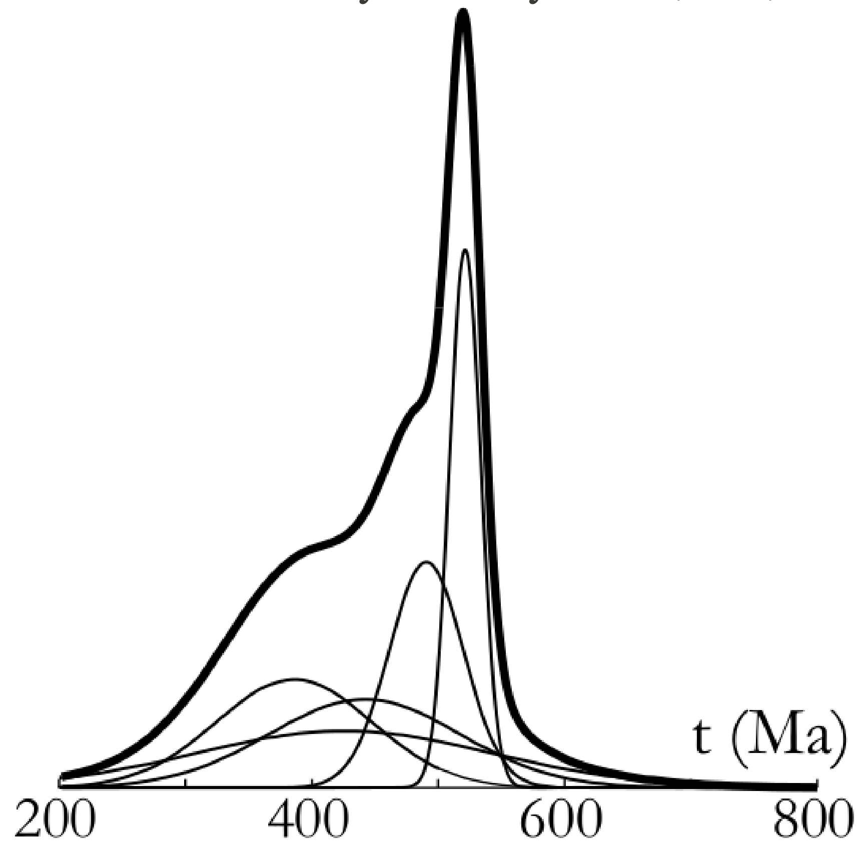


Kernel Density Estimate (KDE)



$$KDE(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right)$$

Probability Density Plot (PDP)



$$PDP(x) = \frac{1}{nh_i}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h_i}\right)$$

A vast body of statistical literature exists about Kernel Density Estimation

Kernel density estimation ×

en.wikipedia.org/wiki/Kernel_density_estimation

Create account   Log in

Article   Talk   Read   Edit   View history   Search

Wikipedia
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox

Print/export
Create a book
Download as PDF
Printable version

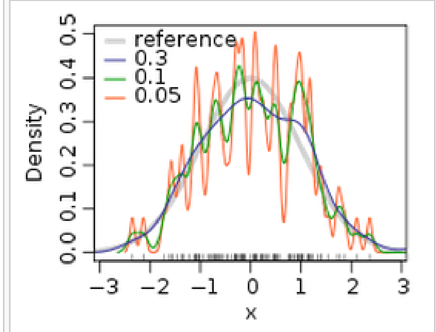
Languages
Deutsch
Français
日本語
Українська
中文

Kernel density estimation

From Wikipedia, the free encyclopedia

It has been suggested that *Multivariate kernel density estimation* be merged into this article or section. (Discuss)
Proposed since September 2010.

In statistics, **kernel density estimation (KDE)** is a non-parametric way to estimate the probability density function of a random variable. Kernel density estimation is a fundamental data smoothing problem where inferences about the population are made, based on a finite data sample. In some fields such as signal processing and econometrics it is also termed the *Parzen–Rosenblatt window* method, after Emanuel Parzen and Murray Rosenblatt, who are usually credited with independently creating it in its current form.[1][2]



Kernel density estimation of 100 normally distributed random numbers using different smoothing bandwidths.

Contents [hide]

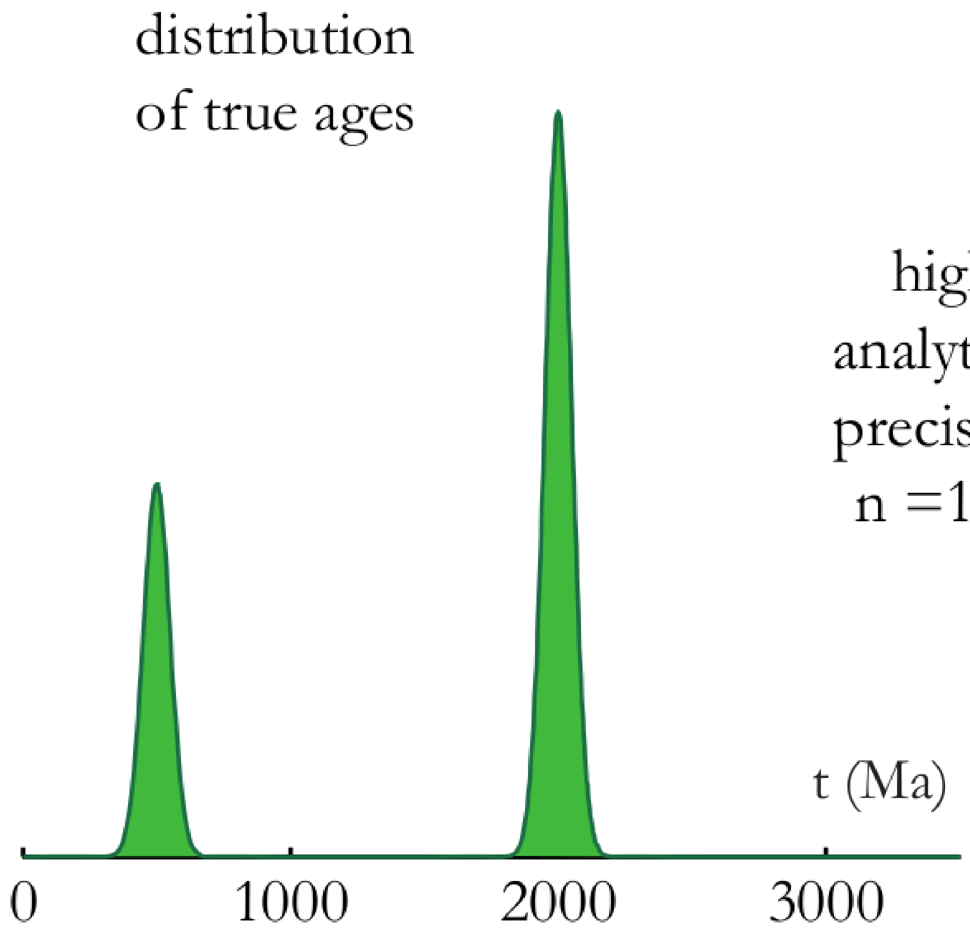
Definition

[edit]

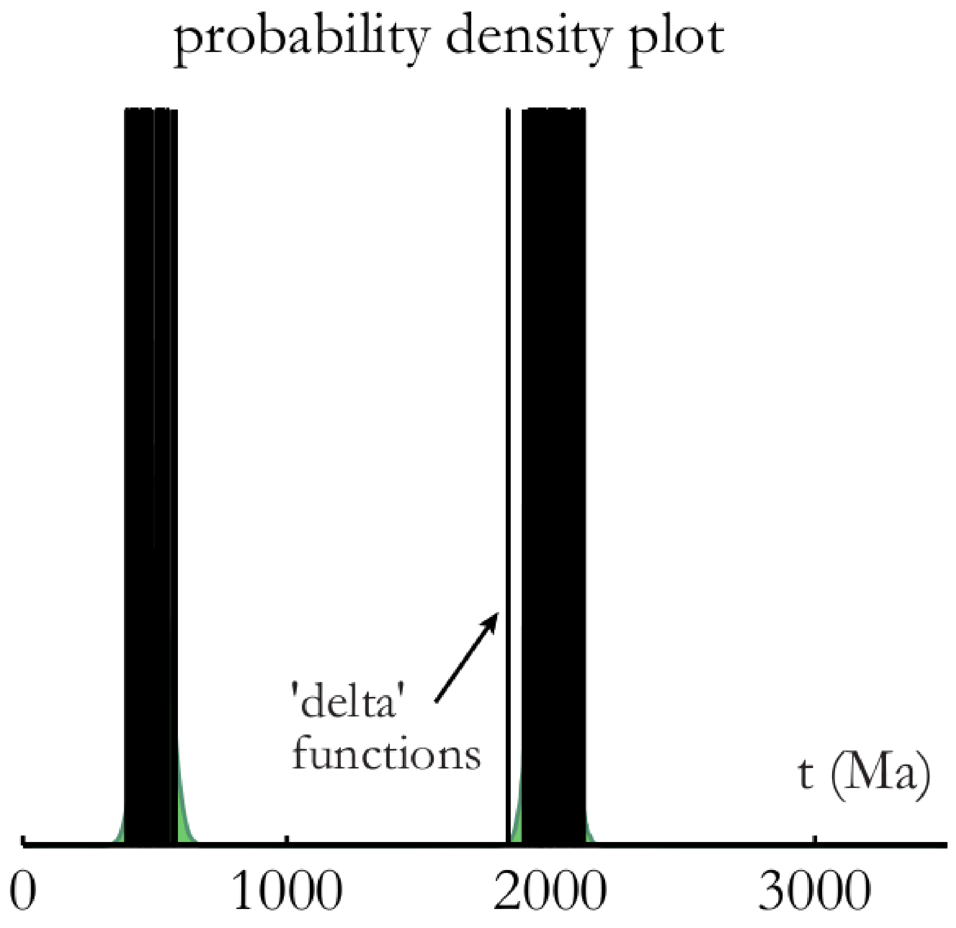
Let $(x_1, x_2, \ldots, x_n)$ be an iid sample drawn from some distribution with an unknown density f. We are interested in estimating the shape of this function f. Its *kernel density estimator* is

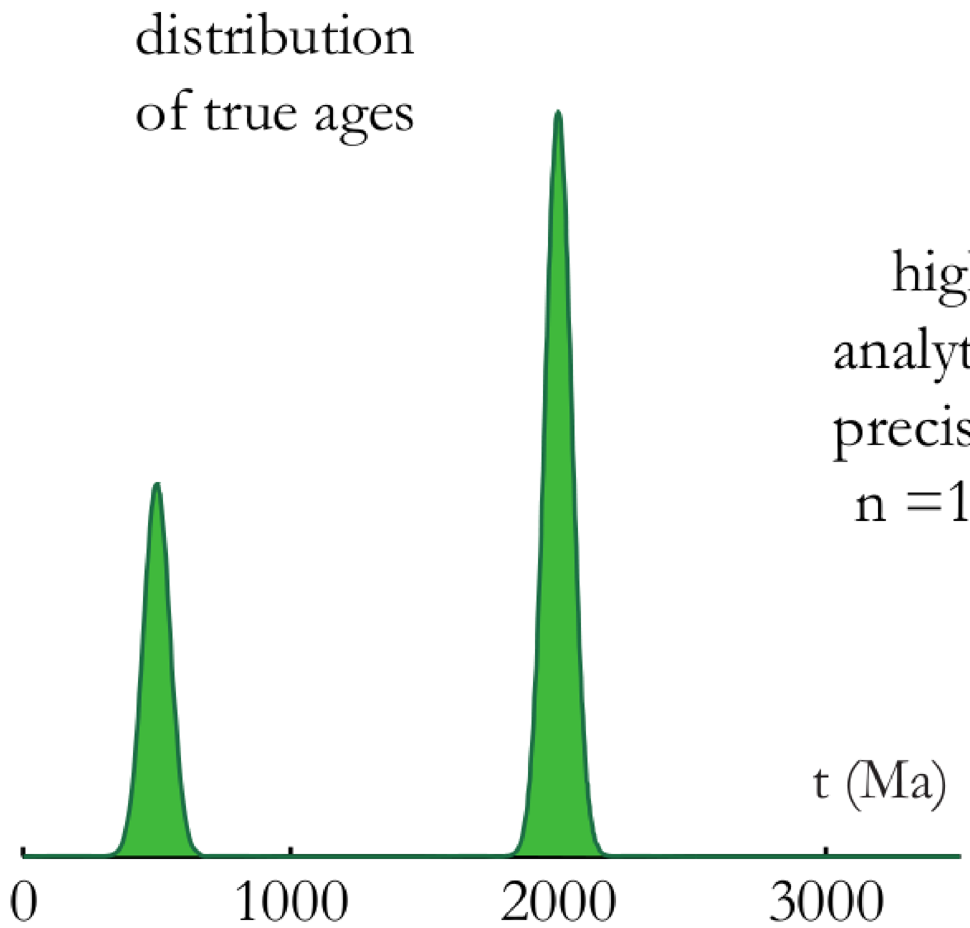
$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) \quad = \frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{x - x_i}{h}\Big),$$

where $K(\bullet)$ is the kernel — a symmetric but not necessarily positive function that integrates to one — and $h > 0$ is a smoothing parameter called the *bandwidth*. A

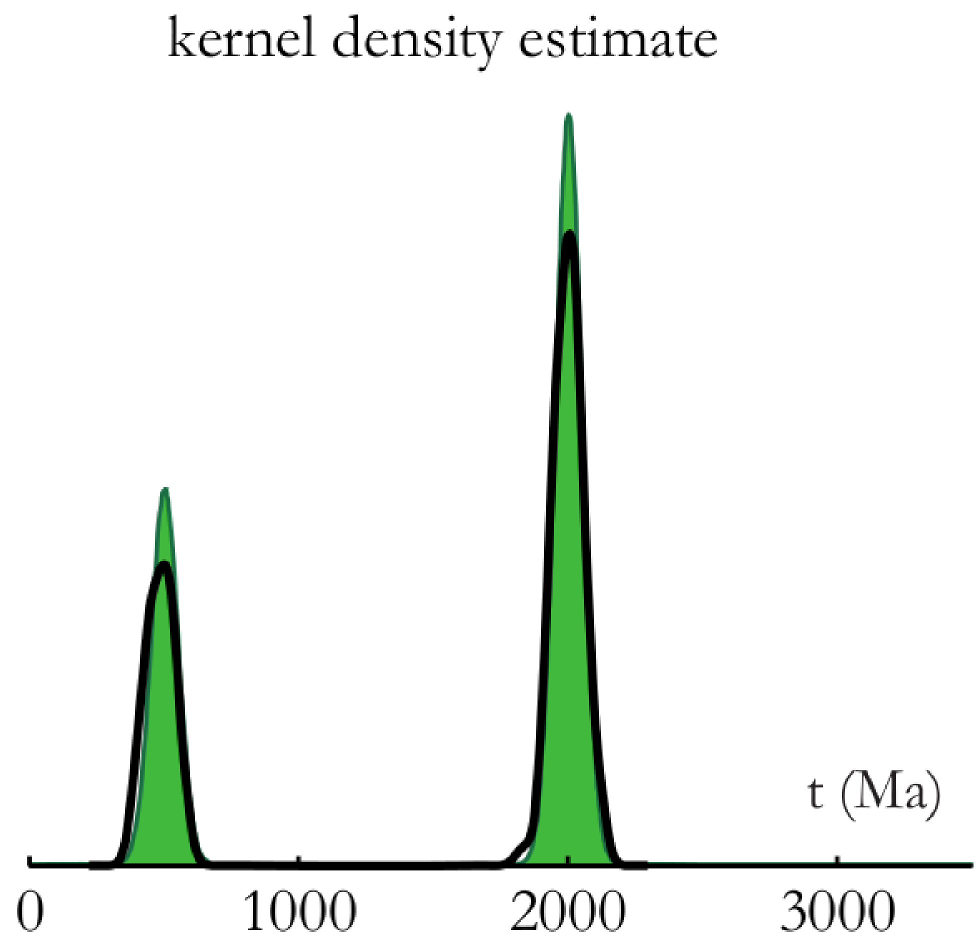


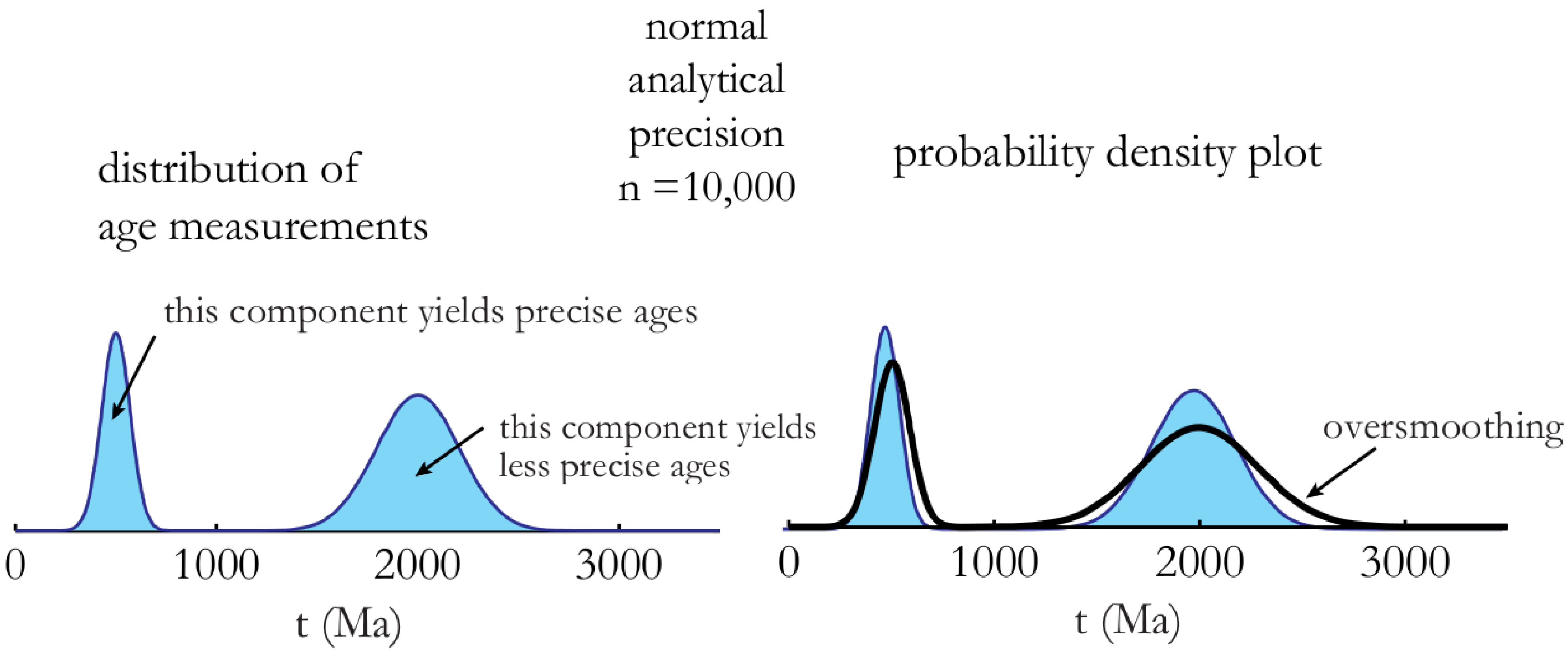
high
analytical
precision
n =117



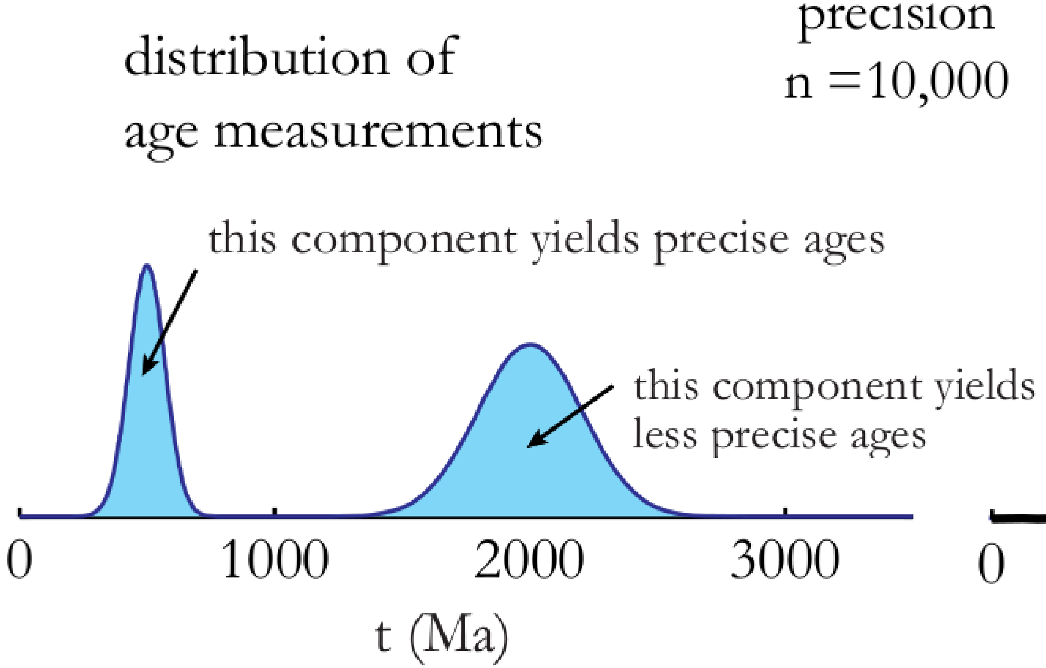


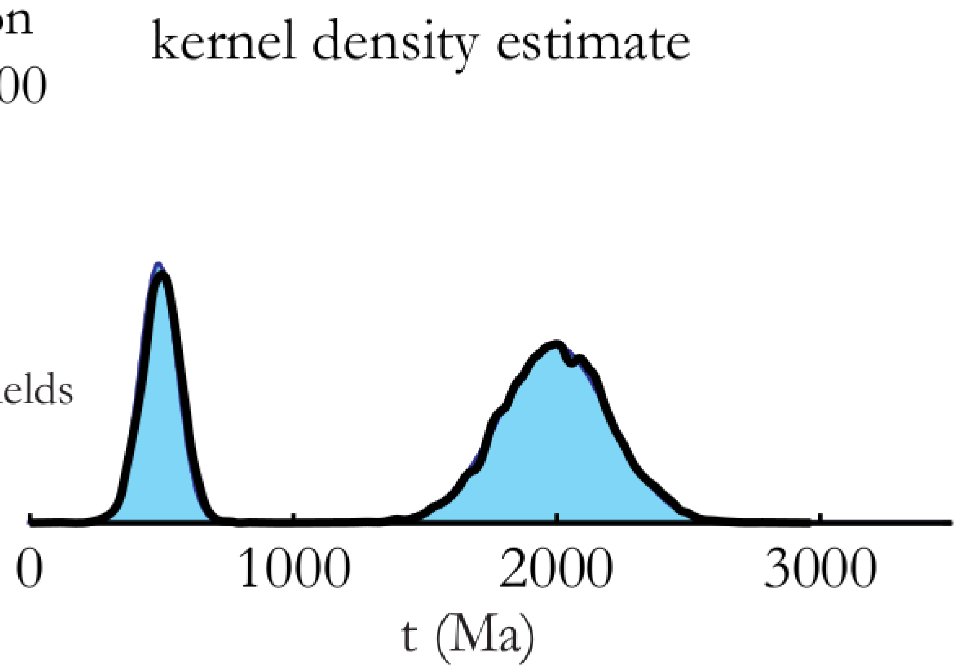
high
analytical
precision
n =117




distribution of
age measurements
normal
analytical
precision
n =10,000
probability density plot
this component yields precise ages
this component yields
less precise ages
oversmoothing
0
1000
2000
3000
t (Ma)
0
1000
2000
3000
t (Ma)

normal
analytical
precision
n =10,000





histograms

$$k = \sqrt{n} \ (Excel)$$

$$k = [\log_2 n + 1] \ (Sturges' \ Rule)$$

$$h = 2 \frac{IQR(x)}{n^{1/3}} \ (Freedman-Diaconis)$$

...

KDEs

$$h = 1.06 \, \sigma \, n^{-1/5} \ (Silverman)$$

$$h_{AMISE} = \frac{R(k)^{1/5}}{m_2(K)^{2/5} R(f'')^{1/5} n^{1/5}}$$

adaptive KDE methods

...

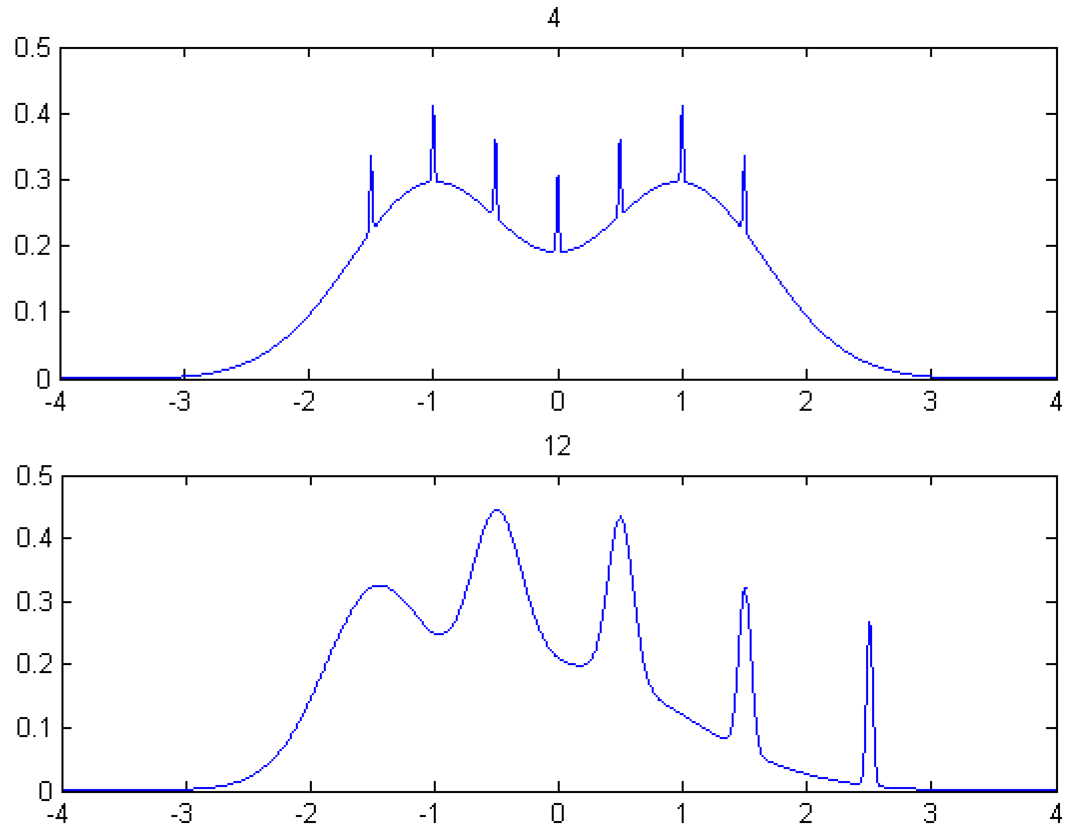
The Annals of Statistics
2010, Vol. 38, No. 5, 2916–2957
DOI: 10.1214/10-AOS799

KERNEL DENSITY ESTIMATION VIA DIFFUSION

BY Z. I. BOTEV[1], J. F. GROTOWSKI AND D. P. KROESE[1]

University of Queensland

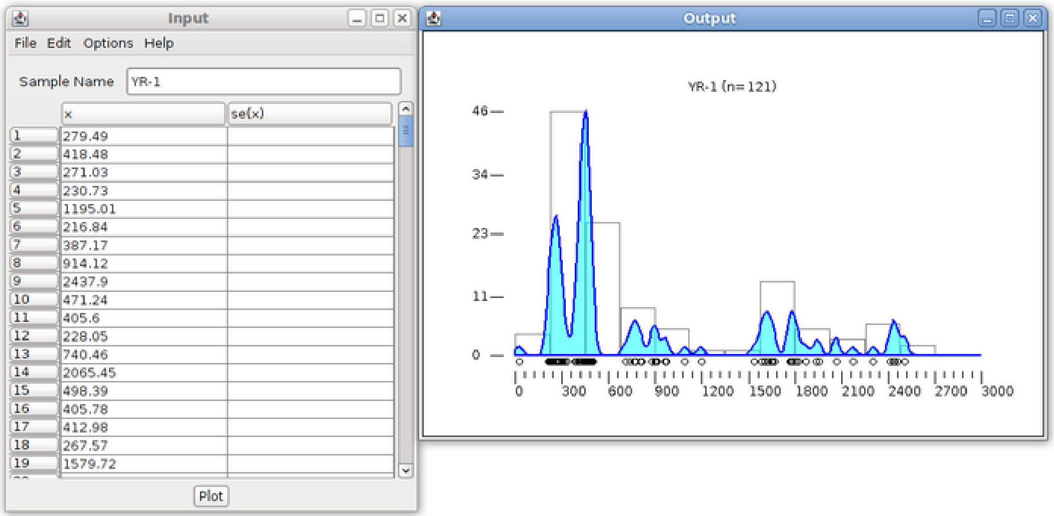
Case	Target density $f(x)$
1	$\frac{1}{2}\mathrm{N}(0,(\frac{1}{10})^2)+\frac{1}{2}\mathrm{N}(5,1)$
2	$\frac{1}{2}\mathrm{N}(0,1)+\sum_{k=0}^{4}\frac{1}{10}\mathrm{N}(\frac{k}{2}-1,(\frac{1}{10})^2)$
3	$\sum_{k=0}^{7}\frac{1}{8}\mathrm{N}(3((\frac{2}{3})^k-1),(\frac{2}{3})^{2k})$
4	$\frac{49}{100}\mathrm{N}(-1,(\frac{2}{3})^2)+\frac{49}{100}\mathrm{N}(1,(\frac{2}{3})^2)+\frac{1}{350}\sum_{k=0}^{6}\mathrm{N}(\frac{k-3}{2},(\frac{1}{100})^2)$
5	$\frac{2}{7}\sum_{k=0}^{2}\mathrm{N}(\frac{12k-15}{7},(\frac{2}{7})^2)+\frac{1}{21}\sum_{k=8}^{10}\mathrm{N}(\frac{2k}{7},(\frac{1}{21})^2)$
6	$\frac{46}{100}\sum_{k=0}^{1}\mathrm{N}(2k-1,(\frac{2}{3})^2)+\sum_{k=1}^{3}\frac{1}{300}\mathrm{N}(-\frac{k}{2},(\frac{1}{100})^2)$ $+\sum_{k=1}^{3}\frac{7}{300}\mathrm{N}(\frac{k}{2},(\frac{7}{100})^2)$
7	$\frac{1}{2}\mathrm{N}(-2,\frac{1}{4})+\frac{1}{2}\mathrm{N}(2,\frac{1}{4})$
8	$\frac{3}{4}\mathrm{N}(0,1)+\frac{1}{4}\mathrm{N}(\frac{3}{2},(\frac{1}{3})^2)$
9	Log-Normal with $\mu=0$ and $\sigma=1$
10	$\frac{1}{2}\mathrm{N}(0,1)+\sum_{k=-2}^{2}\frac{2^{1-k}}{31}\mathrm{N}(k+\frac{1}{2},(\frac{2^{-k}}{10})^2)$



DensityPlotter

www.ucl.ac.uk/~ucfbpve/densityplotter/

DensityPlotter - a Java application for Kernel Density Estimation



downloads	
DensityPlotter.jar Version 2.0 of the program earlier versions	Example input files FTdata.csv: Fission Tracks UPbdata.csv: U-Pb ages

`DensityPlotter` produces publication-ready kernel density estimates, probability density plots, histograms, radial plots and mixture models of (detrital) age distributions. The program is based on, and in fact offers exactly the same functionality as RadialPlotter albeit with a different set of pre-loaded preferences.

Citeable reference:
Vermeesch, P., 2012. On the visualisation of detrital age distributions. Chemical Geology, v.312-313, 19 0

Any feedback is appreciated!