Geological Survey
of Canada

RADIOGENIC AGE AND ISOTOPIC STUDIES: REPORT 13

Current Research 2000-F2

The usefulness and limitations of binned frequency histograms and probability density distributions for displaying absolute age data

K. Sircombe

2000

Natural Resources Canada Ressources naturelles Canada

Canada

Catalogue No. M44-2000/F2E-IN
ISBN 0-660-18228-9

Available in Canada from the
Geological Survey of Canada Bookstore website at:
http://www.nrcan.gc.ca/gsc/bookstore (Toll-free: 1-888-252-4301)

A copy of this publication is also available for reference by depository libraries across Canada through access to the Depository Services Program's website at http://dsp-psd.pwgsc.gc.ca

Price subject to change without notice

Author's address

K. Sircombe *(ksircomb@nrcan.gc.ca)*
Geological Survey of Canada
601 Booth Street
Ottawa, Ontario K1A 0E8

The usefulness and limitations of binned frequency histograms and probability density distributions for displaying absolute age data

K. Sircombe
GSC Ottawa, Ottawa

Sircombe, K., 2000: The usefulness and limitations of binned frequency histograms and probability density distributions for displaying absolute age data; Geological Survey of Canada, Current Research 2000-F2; Radiogenic Age and Isotopic Studies: Report 13, 11 p. (online; http://www.nrcan.gc.ca/bookstore)

Abstract: Initial assessment and visual communication of the salient features of large sets of geochronological age data are commonly achieved with binned frequency histograms or probability density distributions. Both are estimates of the sample distribution and each has inherent limitations. Although simple and effective at conveying frequency information, histograms have two important limitations, i.e. analytical errors are discarded and diagram appearance is potentially vulnerable to bias because of arbitrary decisions about interval widths. A method for assessing the efficiency of bin widths is presented. Age probability density distribution diagrams use a variable gaussian kernel method that accounts for the analytical error of individual datum. This also provides for the standardization of display by avoiding arbitrary bin width selection. However, these diagrams are limited by the lack of visual frequency data. Thus to maximize displayed information in some cases it may be beneficial to combine both methods.

Résumé : Les histogrammes de la fréquence par casiers ou les distributions de densité de probabilité permettent communément d'effectuer une évaluation initiale et une communication visuelle des caractéristiques principales de grands ensembles de données géochronologiques. Dans les deux cas, ce sont des estimations limitées de la répartition des échantillons. Bien que simples et efficaces pour fournir des informations sur la fréquence, les histogrammes comportent des limitations importantes : les erreurs analytiques sont éliminées et l'apparence des diagrammes est susceptible d'être biaisée par des décisions arbitraires sur la largeur des casiers. On présente ici une méthode pour évaluer l'efficacité de la la largeur des casiers. Les diagrammes de distribution de la densité de probabilité des âges utilisent une méthode à noyau gaussien variable qui tient compte de l'erreur analytique de chaque donnée. On peut ainsi uniformiser l'affichage en évitant un choix arbitraire de la largeur des casiers. Cependant, l'absence de données visuelles sur la fréquence limitent l'utilisation de ces diagrammes. Pour maximiser les données montrées, il pourrait être préférable, dans certains cas, de combiner les deux méthodes.

INTRODUCTION

The acquisition of relatively large sets of U-Pb isotopic age data has become routine in secondary ionization mass spectrometry (e.g. Dodson et al., 1988; Whitehouse et al., 1997; Rainbird et al., 1998; Geslin et al., 1999), and in isotope dilution thermal ionization mass spectrometry (e.g. Davis et al., 1994; Gehrels and Dickinson, 1995). Uranium-lead isotopic data (i.e. ^{206}Pb-^{238}U and ^{207}Pb-^{235}U) are typically displayed using either a Wetherill- or Tera-Wasserburg-style bivariate concordia plot (e.g. Fig. 1; Wetherill, 1963; Tera and Wasserburg, 1972, 1974). The primary purpose of these diagrams is to convey information about a set of data that is salient to the analytical process: accuracy (relationship with concordia and repeat analyses), precision (analytical error displayed as error ellipses or equivalent), and sample size. However, as sample size increases, concordia diagrams can quickly become visually cluttered, making data assessment for modes, ranges and proportions difficult. Also, concordia diagrams are potentially meaningless for nonspecialist and/or nontechnical audiences. Therefore it is often advantageous to display the data in terms of the univariate absolute age calculated from one of the isotopic ratios.

The most common alternative forms of display are the binned frequency histogram and probability density distribution. (Fission-track derived age data is often displayed using radial age plots (Galbraith, 1989). These plots provide a lot of information about the age data including analytical error, but can be difficult for a non-specialist to interpret and will not be discussed here). The use of such diagrams is most common in sedimentary provenance studies involving relatively large sets of age data (e.g. Morton et al., 1996; Pell et al., 1997; Sircombe, 1999; Sircombe and Freeman, 1999), but is also appropriate for the analysis of metamorphic and/or igneous rocks where the display of complex age patterns is required (e.g. Harley and Black, 1997). This paper describes the mathematical basis of traditional binned frequency histograms and probability density distributions and then compares and contrasts the advantages and limitations of both methods for displaying absolute age data.

BRIEF NOTE ON SCREENING ANALYSES PRIOR TO DISPLAY

Displaying univariate age data means that the visual information regarding accuracy seen in a concordia diagram is lost. Therefore, prior to generating diagrams of age data, the data are usually systematically filtered for accuracy, particularly in sensitive high-resolution ion microprobe (SHRIMP) studies. This is usually judged by an arbitrary constraint on concordance as defined by the $^{206}Pb/^{238}U$ age ratioed to the $^{207}Pb/^{206}Pb$ age and multiplied by 100 to produce a percentage value. A typical concordance constraint for detrital zircon analyses may be between 95% and 105% and is particularly applied in material older than ~1500 Ma in SHRIMP studies. For younger material examined by SHRIMP analysis, the ^{207}Pb-^{235}U ratio and thus the ^{207}Pb-^{206}Pb ratio and calculated ages become increasingly imprecise due to poorer counting statistics. In those cases, the value of a concordance constraint and the accuracy filtering methodology has to be judged on a case-by-case basis.

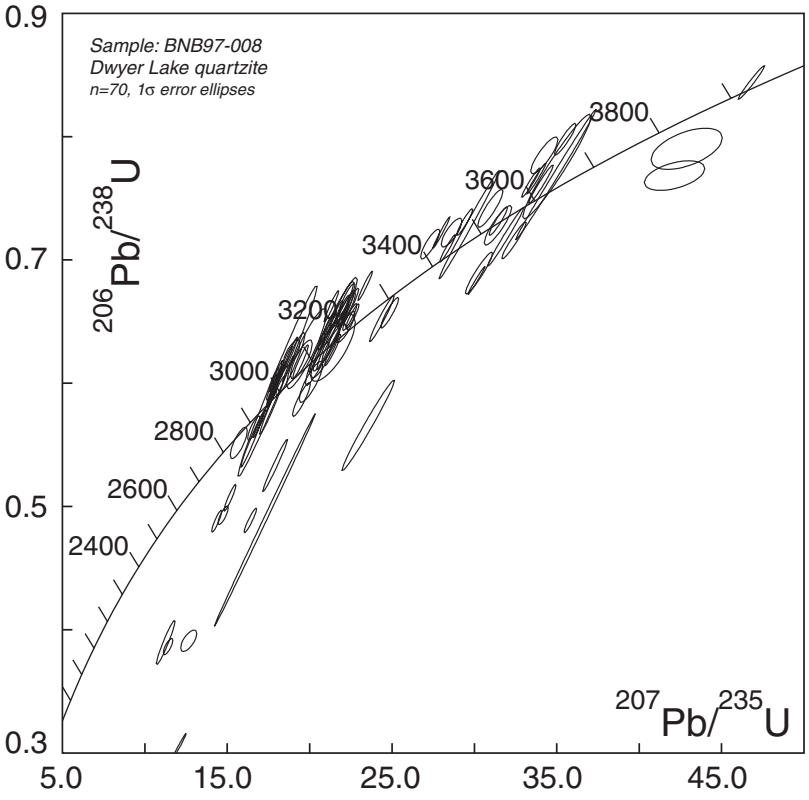

Figure 1.

Example of a set of detrital zircon age data displayed using a Wetherill style concordia plot (Dwyer Lake quartzite, Slave Province, Northwest Territories; diagram after *Bleeker et al., 2000).*

Filtering of data raises an important issue for both binned frequency histograms and age probability density distributions. If the purpose of the display is to avoid the visual clutter of a concordia diagram, then a filtering process will remove those discordant results that could contain significant geological information. Care should be taken to be clear about what sort of age data is being displayed and why. In some cases it may be beneficial to include both concordant and discordant results within the age display diagram (e.g. Roback and Walker, 1995, Fig. 8; Morton et al., 1996, Fig. 3) to illustrate that data considered discordant exist and have been considered. These decisions should be on a individual basis depending on the purpose of data display, but at the very minimum the filtering method, if any, and the type of data displayed should be clearly stated.

MATHEMATICAL BACKGROUND AND LIMITATIONS OF BINNED FREQUENCY HISTOGRAMS

Traditionally, relative large numbers of absolute age data have been displayed using a binned frequency histogram. In many cases such a histogram is an important means for analyzing and communicating salient features of a set of age data and thus the geological interpretation of that data. In particular, modes, ranges, and proportions displayed in a binned frequency histogram may relate to the timing, duration, and relative significance of geological events. Because the binned frequency histograms provide an important step in the graphical analysis and communication of age data, it is equally important that the analyst be aware of the definitions and inherent limitations of the method.

With a set of age data, typically measured in Ma, $\{x_1, ..., x_N\}$, the histogram is generated by dividing the age range of interest, anchored at an origin x_0, into a set of K equal-sized bins, each with a bin width h, defined by the bin limits $\{x_0 + jh, x_0 + (j + 1)h\}$ for positive integer values of j up to K (Silverman, 1986; Scott, 1992; Simonoff and Udina, 1997). (Alternatively, bins may also be termed *class intervals* as defined between *class limits*.) An important mathematical definition regarding histograms that is often overlooked must be made at this point. A histogram is an estimate of the distribution of a sample of univariate data or $f(x)$. In the simplest case, the *frequency histogram*, the estimate is as follows:

$$\hat{f}(\chi) = n_j, (x_A + jh \le x_i < x_A + (j+C)h) \qquad (1)$$

where n_j is the number of age data, x_i, within the jth bin $(x_0 + jh, x_0 + (j + 1)h)$. A *relative frequency* histogram is given by the ratio of the count in each bin against the total N as follows:

$$\hat{f}(x) = \frac{n_j}{N}, (x_A + jh \le x_i < x_A + (j+C)h) \qquad (2)$$

Both are plotted using the frequency or relative frequency values to define the height of columns. At this stage, a distinction between histograms used for data presentation and for density estimation can be made. In the frequency or relative frequency histogram, the bin counts and thus column size can be used to convey area as a density estimate of the sample distribution. Alternatively, the bin counts could be displayed directly or as a stem-and-leaf plot (Tukey, 1977), effectively rendering the histogram in symbols rather than graphically. The formal definition of a histogram used for *density estimation* is as follows:

$$\hat{f}(x) = \frac{n_j}{Nh}, (x_A + jh \le x_i < x_A + (j+C)h) \qquad (3)$$

(Silverman, 1986; Scott, 1992; Simonoff and Udina, 1997), where the count is also ratioed to the bin width h; thus the bin width on the plot is an integral aspect of the display and in effect it is the *area* of the column that conveys information, not just the height of the column representing the bin. Although such a density estimation histogram described in equation (3) is not typically used for the display of age data, a knowledge of the distinction between the display of frequency (column height) and density estimation (column area) is important.

The methodology of binned frequency histograms involves two critical limitations that may diminish its usefulness as a means of displaying age data if more detailed communication is required. Firstly, the analytical error $\{e_1, ..., e_N\}$, which is an intrinsic part of any good geochronological analysis, is not considered in constructing a binned frequency histogram as described above. The sample distribution is therefore estimated on the basis of the age data alone, and precision, a salient feature of the analytical process, is discarded. If the age data with analytical errors are accessible, this estimation in the form of a frequency histogram may not cause a problem. In cases where a diagram displaying age data is the principal method for communicating results, discarding the errors may lead to misinterpretation. In a binned frequency histogram, an analysis with a relatively large analytical error will be treated in the same way as an analysis with a relatively smaller error, even though the two may not be strictly comparable. For instance, as shown in Figure 2, two analyses have the same age value of 2015 Ma, but one has an analytical error of 20 Ma and the other, 2 Ma (both at 1 s.e.). If the bin limits are 2000 Ma and 2020 Ma and assuming a Gaussian

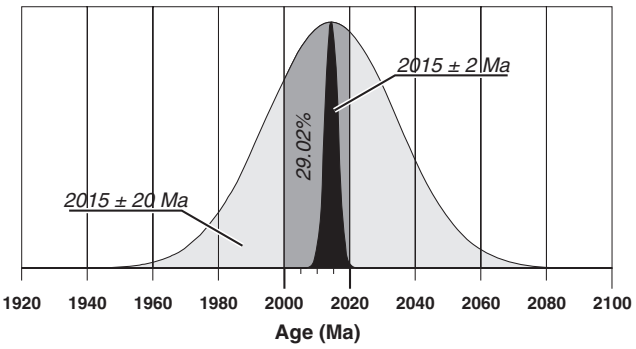

Figure 2. *Example of two age estimates of the same mean, but different variance, representing age data cut by a histogram bin. This figure illustrates that depending on analytical error inherent in the age data and histogram bin width, the bin may only represent a proportion of the age estimate.*

Filtering of data raises an important issue for both binned frequency histograms and age probability density distributions. If the purpose of the display is to avoid the visual clutter of a concordia diagram, then a filtering process will remove those discordant results that could contain significant geological information. Care should be taken to be clear about what sort of age data is being displayed and why. In some cases it may be beneficial to include both concordant and discordant results within the age display diagram (e.g. Roback and Walker, 1995, Fig. 8; Morton et al., 1996, Fig. 3) to illustrate that data considered discordant exist and have been considered. These decisions should be on a individual basis depending on the purpose of data display, but at the very minimum the filtering method, if any, and the type of data displayed should be clearly stated.

MATHEMATICAL BACKGROUND AND LIMITATIONS OF BINNED FREQUENCY HISTOGRAMS

Traditionally, relative large numbers of absolute age data have been displayed using a binned frequency histogram. In many cases such a histogram is an important means for analyzing and communicating salient features of a set of age data and thus the geological interpretation of that data. In particular, modes, ranges, and proportions displayed in a binned frequency histogram may relate to the timing, duration, and relative significance of geological events. Because the binned frequency histograms provide an important step in the graphical analysis and communication of age data, it is equally important that the analyst be aware of the definitions and inherent limitations of the method.

With a set of age data, typically measured in Ma, $\{x_1, ..., x_N\}$, the histogram is generated by dividing the age range of interest, anchored at an origin x_0, into a set of K equal-sized bins, each with a bin width h, defined by the bin limits $\{x_0 + jh, x_0 + (j + 1)h\}$ for positive integer values of j up to K (Silverman, 1986; Scott, 1992; Simonoff and Udina, 1997). (Alternatively, bins may also be termed *class intervals* as defined between *class limits*.) An important mathematical definition regarding histograms that is often overlooked must be made at this point. A histogram is an estimate of the distribution of a sample of univariate data or $f(x)$. In the simplest case, the *frequency histogram*, the estimate is as follows:

$$\hat{f}(\chi) = n_j, (x_A + jh \le x_i < x_A + (j+C)h) \qquad (1)$$

where n_j is the number of age data, x_i, within the jth bin $(x_0 + jh, x_0 + (j + 1)h)$. A *relative frequency* histogram is given by the ratio of the count in each bin against the total N as follows:

$$\hat{f}(x) = \frac{n_j}{N}, (x_A + jh \le x_i < x_A + (j+C)h) \qquad (2)$$

Both are plotted using the frequency or relative frequency values to define the height of columns. At this stage, a distinction between histograms used for data presentation and for density estimation can be made. In the frequency or relative frequency histogram, the bin counts and thus column size can be used to convey area as a density estimate of the sample distribution. Alternatively, the bin counts could be displayed directly or as a stem-and-leaf plot (Tukey, 1977), effectively rendering the histogram in symbols rather than graphically. The formal definition of a histogram used for *density estimation* is as follows:

$$\hat{f}(x) = \frac{n_j}{Nh}, (x_A + jh \le x_i < x_A + (j+C)h) \qquad (3)$$

(Silverman, 1986; Scott, 1992; Simonoff and Udina, 1997), where the count is also ratioed to the bin width h; thus the bin width on the plot is an integral aspect of the display and in effect it is the *area* of the column that conveys information, not just the height of the column representing the bin. Although such a density estimation histogram described in equation (3) is not typically used for the display of age data, a knowledge of the distinction between the display of frequency (column height) and density estimation (column area) is important.

The methodology of binned frequency histograms involves two critical limitations that may diminish its usefulness as a means of displaying age data if more detailed communication is required. Firstly, the analytical error $\{e_1, ..., e_N\}$, which is an intrinsic part of any good geochronological analysis, is not considered in constructing a binned frequency histogram as described above. The sample distribution is therefore estimated on the basis of the age data alone, and precision, a salient feature of the analytical process, is discarded. If the age data with analytical errors are accessible, this estimation in the form of a frequency histogram may not cause a problem. In cases where a diagram displaying age data is the principal method for communicating results, discarding the errors may lead to misinterpretation. In a binned frequency histogram, an analysis with a relatively large analytical error will be treated in the same way as an analysis with a relatively smaller error, even though the two may not be strictly comparable. For instance, as shown in Figure 2, two analyses have the same age value of 2015 Ma, but one has an analytical error of 20 Ma and the other, 2 Ma (both at 1 s.e.). If the bin limits are 2000 Ma and 2020 Ma and assuming a Gaussian

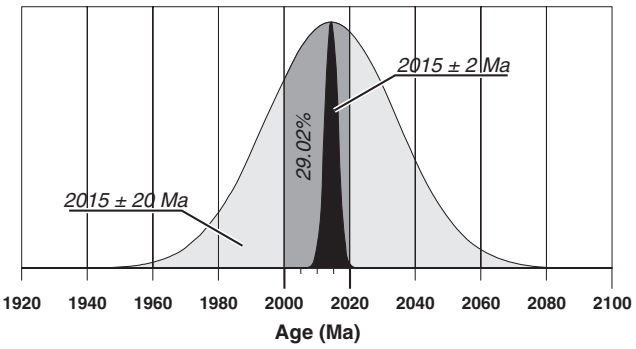

Figure 2. *Example of two age estimates of the same mean, but different variance, representing age data cut by a histogram bin. This figure illustrates that depending on analytical error inherent in the age data and histogram bin width, the bin may only represent a proportion of the age estimate.*

Table 1. Examples of bin width and other details of binned frequency histogram displays from a variety of references. Mean error and efficiency (*E*, explained in text) calculated for a subset of presented data and intended for only for broad indicative purposes.

Bin width	Range	Bin width as %Range	Mean error	Efficiency (*E*)	Reference
5 Ma	70 Ma	7.14%	2 Ma[1]	70%	Davis et al., 1994, Fig. 8B
20 Ma	3400 Ma	0.588%	5 Ma[1]	80%	Gehrels and Dickinson, 1995, Fig. 6
25 Ma	4000 Ma	0.625%	20 Ma[2]	60%	Morton et al., 1996, Fig. 3
33.33 Ma	1900 Ma	1.75%	50 Ma[3]	30%	Scott and Gauthier, 1996, Fig. 4
100 Ma	2800 Ma	3.57%	3 Ma[1]	98%	Roback and Walker, 1995, Fig. 8 †

[1]: thermal ionization mass-spectrometer analyses
[2]: SHRIMP analyses
[3]: laser-ablation microprobe inductively coupled mass spectrometer analyses
†: Roback and Walker (1995) histograms compiled from data in Ross et al. (1991, 1992).

distribution for the age estimate, only 29.02% of the distribution defined by the first analysis (2015 ± 20 Ma) lies within the bin limits. (For the purposes of discussion here, the distribution associated with an individual age calculation will be referred to as the 'age estimate' and defined as a Gaussian distribution defined by a mean x_i (calculated age) and standard deviation e_i (analytical error). In the strictest sense, the assumption of a symmetrical Gaussian distribution is not valid for absolute age calculations based on U-Pb isotopic measurements. Because the relationship between measured U-Pb and Pb-Pb ratios and absolute age is not linear, analytical error propagation results in a slightly asymmetrical distribution, particularly those with relatively large analytical errors. However, for the sake of discussion here, it is assumed that the asymmetrical nature of age estimates is only rarely significant.) In effect this is saying that although the mean of the age estimate distribution is 2015 Ma and thet value lies within the bin limits 2000 and 2020 Ma, there is a 70.98% probability that the 'true' value of the age lies outside the bin limits. In comparison, 99.38% of the age estimate defined by the second analysis (2015 ± 2 Ma) lies within the bin limits, so the bin may be considered representative. However, if the bin width was the same, but the limits were, for example, 1996 Ma and 2016 Ma, then even the relatively more precise second analysis would have 30.85% of its age estimate outside the bin limits.

From this example it can be seen that the estimate of the sample distribution made by the binned frequency histogram only counts the mean of the age estimates defined by the sample ages and their errors. Potentially a lot of information about the precision of individual measurements and the overall sample distribution may be lost. This example focuses attention on the second limitation inherent to binned histograms, i.e. the size and location of the bins themselves. A histogram's appearance, and thus its potential interpretation, is a balance between too much detail with narrow bin widths (undersmoothing) or too little detail with wide bin widths (oversmoothing). In various published examples (Table 1), the choice of bin width for displaying age data varies greatly from 5 Ma (Davis et al., 1994, Fig. 8B) through 20 Ma (Gehrels and Dickinson, 1995, Fig. 6), and from 33.33 Ma (Scott and Gauthier, 1996, Fig. 4) to 100 Ma (Roback and Walker, 1995, Fig. 8, using data from Ross et al., 1991, 1992). Using SHRIMP-derived data, Morton et al. (1996, p. 917) defined bin widths of 25 Ma on the basis of the precision of the analyses, although no explanation is given of how this was done.

Because bin width can be an arbitrary decision, accidental or even deliberate bias in the appearance and interpretation of the binned frequency histogram is possible. If the full set of data is accessible and other statistical analyses are done to describe the sample distribution, this may not present a problem. Nevertheless, in a case where a binned frequency histogram is used to quickly convey salient information without statistical details, the potential for misinterpreting subtleties exists. Figure 3 uses a set of example age data (Dwyer Lake quartzite, Slave Province, Northwest Territories; Bleeker et al., 2000; K. Sircombe, W. Bleeker, and R. Stern, work in progress, 2000) to illustrate the change in appearance of the binned frequency histogram as the bin width changes. Although the bimodal (~2850 Ma and ~3150 Ma) and dispersive nature of the sample distribution is apparent in all the histograms, some subtleties can also be seen. For instance, the relative prominence of the two modes changes, particularly between bin widths 33.33 Ma (Fig. 3c) and 50 Ma (Fig. 3d) where the ~3150 Ma mode appears to become dominant.

A number of methods exist for calculating the optimal bin width for a given set of data. For instance, Doane (1985) proposed the following modification to the earlier Sturges (1926) method:

$$\hat{h} = \frac{\max\{x_c, \ldots, x_N\} - \min\{x_c, \ldots, x_N\}}{1 + \log_2 N + K_e} \qquad (4)$$

where $\hat{h}$ is the estimate of the optimal bin width, N is the sample size, and K_e is the extra classes proposed by Doane (1985) to account for potential skewness in the data, as follows:

$$K_e = \log_2\left(1 + \frac{\sqrt{b_1}}{\sigma\sqrt{b_1}}\right) \qquad (5)$$

where $\sqrt{b_1}$ is a measure of skewness,

$$\sqrt{b_1} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^3}{\left[\sum_{i=1}^{N}(x_i - \bar{x}^2)\right]^{3/2}} \quad (6)$$

and $\sigma\sqrt{b_1}$ is defined by the sample size,

$$\sigma\sqrt{b_1} = \sqrt{\frac{6(N-2)}{(N+1)(N+3)}} \quad (7)$$

Many statistical packages use this rule or a modification of it to calculate the default bin width for a set of data (Wand, 1996). Recognizing that Sturges' (1926) rule and modifications may lead to oversmoothing, more recent formulations include Scott's (1979) normal reference rule,

$$\hat{h} = 3.49\,\hat{\sigma}\,n^{-1/3} \quad (8)$$

where $\hat{\sigma}$ is an estimate of the standard deviation, typically the minimum of either the sample standard deviation or the interquartile range (Wand, 1996). The interquartile range is equal to Q_3 (upper quartile, or 75% percentile) subtract Q_1 (lower quartile, or 25% percentile). Increasingly sophisticated calculations are based on minimizing the difference between the estimate represented by the histogram and the actual sample distribution (Scott, 1992; Wand, 1996; Simonoff and Udina, 1997). For the two sets of example data

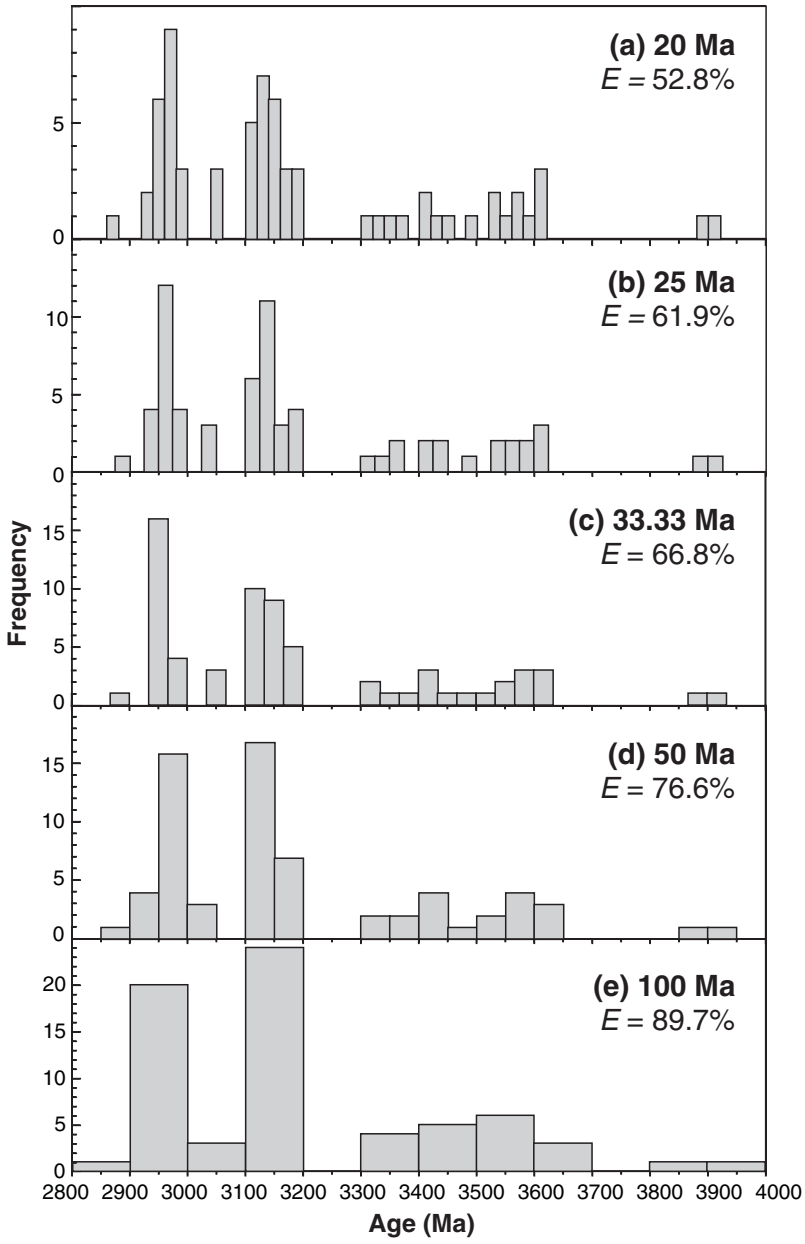

Figure 3.

Dwyer Lake quartzite detrital zircon concordant (95–105%) age data displayed using a series of binned frequency histograms with varied bin widths cited in the literature to demonstrate how the appearance and thus the potential interpretation can vary according to an arbitrary decision about bin width. Mean error of age data = 15.8 Ma. E= calculated efficiency of the bin width at capturing the individual age estimates.

discussed here, equations (4) and (8) yield optimal bin widths of 136 Ma and 210 Ma for the Dwyer Lake sample. In comparison, another example set of detrital zircon age data (George Lake metagreywacke, Slave Province, Northwest Territories; K. Sircombe, W. Bleeker, and R. Stern, work in progress, 2000) yields optimal bin widths of 36 Ma and 22 Ma for equations (4) and (8). This illustrates that although a useful starting point in visualizing an individual set of data, such optimal binning methods do not necessarily produce a standard bin width for easy visual comparison of different sets of data. Critically, because these calculations assume an underlying Gaussian distribution in the sample data, they may not be applicable to sets of data that have strongly non-Gaussian distributions (Scott, 1979, 1992). This may be the case particularly in detrital age data where the sample distribution may be multimodal, dispersed, and complex, like the Dwyer Lake sample.

Another approach to calculating bin width is briefly discussed here for the purpose of completeness. Using information theory, bin widths can be defined using the concept of maximum entropy (Full et al., 1984). Basically, entropy can be applied as a measure of the contrast between bins in a histogram. The greater the calculated entropy, the greater the amount of information conveyed by the bin structure. Histograms with low entropy have a large difference between bins, with the lowest state occurring when all the values fall within a single bin. Histograms with high entropy have little difference between bins. Maximum entropy can be obtained if the bin widths are made unequal and arranged so that the analyses are evenly distributed among the intervals. The optimum number of bins for a set of sample data can also be derived using information theory (Torley, 1998), but the binning results depend on the set of samples being analyzed and thus are not universally applicable for comparison. Therefore, although useful for further mathematical examination of a set of age data, such a histogram will generally be unacceptable for displaying age data for the purpose of geological interpretation. The use of the information theory method also does not address the limitation regarding analytical errors.

To define an appropriate bin width for the age data described here, an empirical approach has been taken to assess the efficiency of a variety of histogram bin widths to represent the individual age estimates in a sample. A set of 100 random ages were generated from a uniform distribution and assigned random errors from a normal distribution with a range of means {1, 2, 5, 10, 20, 30} Ma. It was found that the variance in the distributions from which these random errors were selected did not, at least to this level of analysis, effect the results. Two sets of real data, for the Dwyer Lake quartzite and the George Lake metagreywacke, were also examined and have mean errors of 15.8 and 10.8 Ma at 1 s.e. respectively. Both can be considered typical of mean errors in SHRIMP detrital analysis. For each individual age estimate in a sample, the proportion of the age estimate that is captured within the same bin as the age is calculated (as illustrated in Figure 2). The mean of these values for each set of data is taken as a proxy for the efficiency (E) of the bin width in estimating the age data as follows:

$$E = \frac{\sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}} \int_{z_i^L}^{z_i^U} exp^{-x^2/2}\, dx}{N} \qquad (9)$$

which involves summing the normal distribution function for each sample x_i bounded by standard normal variables defined on the basis of the lower and upper limits of bin j about x_i,

$$z_i^L = \frac{(x_o + jh) - x_i}{\sigma_i}$$
$$z_i^U = \frac{(x_o + (j+1)h) - x_i}{\sigma_i} \qquad (10)$$

Figure 4 shows that for a set of age data with a mean error of 1 Ma (more typical of thermal ionization mass spectrometer analysis) that a bin width of 10 Ma is enough to capture 90% of the age estimates. Table 1 also includes an assessment of the mean error and calculated efficiency for the literature histograms discussed above with a wide range of results.

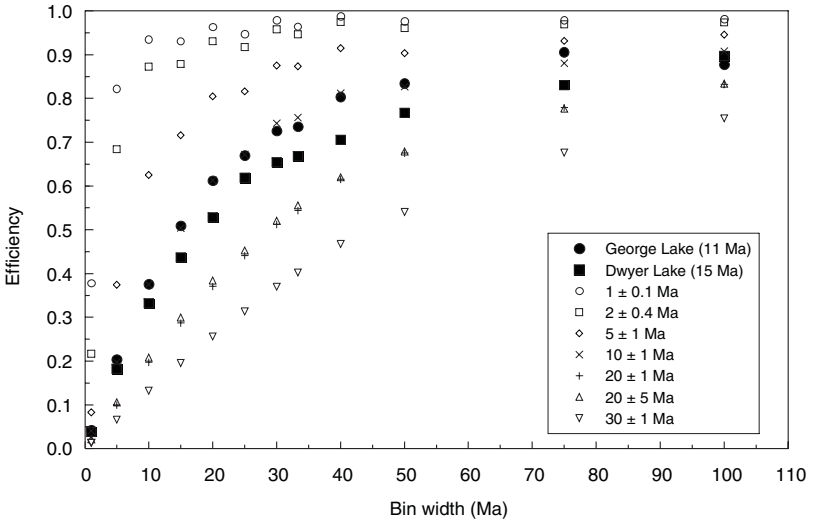

Figure 4.

Relationship between bin width and bin width efficiency (mean proportion of age estimates within same bin as age) for a variety of real and simulated age data with mean errors ranging from 1 Ma to 30 Ma.

For the Dwyer Lake and George Lake samples, a 10 Ma bin width will capture only about 30% of the age estimates; a 20 Ma bin width is needed to go beyond 50% efficiency, but 90% efficiency is not approached until bin widths of over 75 Ma are used. Using this approach as a basis for quantitatively assessing the suitability of the chosen histogram bin width, presented histograms should include a description of the calculated efficiency of the bin width selected for that data. Table 2 lists some suggested bin widths to achieve 50% and 90% efficiency for age data with a variety of mean errors based on the empirical calculations discussed above, although calculations based on the actual data being displayed are recommended. The mean of the sample errors should also be included in the diagram description. In some cases where the range of errors is extremely skewed, it may be preferably to report the median of the sample errors.

Finally, to complete the discussion of parameters affecting histograms, the selection of an origin in a binned histogram display can also significantly alter the appearance and potential interpretation of the data (Simonoff and Udina, 1997). However, in the case of age data presentation, this effect should be minimal because for the general aesthetic sense of 'rounded' limits proposed by Doane (1985), the origin should either be 0 Ma or an integer multiple of the bin width. For example, with a bin width of 25 Ma and minimum value at 2037 Ma, the histogram origin would not be 2036 Ma, but 2000 Ma or 2025 Ma. Any selection of a less orthodox origin would require detailed justification in terms of potential instability in the appearance of the histogram (Simonoff and Udina, 1997).

Table 2. Suggested bin widths required to reach 50% and 90% efficiency in a set of age data with various mean errors based on empirical analysis described in text.

Mean error	Bin width required for	
	>50% efficiency	>90% efficiency
1 Ma	2 Ma	10 Ma
2 Ma	5 Ma	20 Ma
5 Ma	10 Ma	40 Ma
10 Ma	15 Ma	100 Ma
15 Ma	20 Ma	120 Ma
20 Ma	30 Ma	160 Ma
30 Ma	45 Ma	220 Ma
50 Ma	75 Ma	360 Ma

CONSTRUCTING AND INTERPRETING AGE PROBABILITY DENSITY DISTRIBUTIONS

The probability density distribution (PDD) diagram (e.g. Figure 5; Dodson et al., 1988) is a graphical approach to the display of age data that attempts to address the limitations of binned frequency histograms. (The original probability plots applied to detrital zircon age data were presented in Dodson et al. (1988) and were the product of a technique and program ("Nouveau Stats") developed by Dr. P. Zeitler then at the Research School of Earth Sciences of the Australian National University (I.S. Williams, written comment.) Technically, the age probability density distribution is another estimate of the sample distribution like the binned frequency histogram. In this case, it produces a density estimate of the sample distribution using a Gaussian kernel (Silverman, 1986) that varies with each individual age estimate. The shape of these age estimates, and thus the kernel, will vary from a narrow, tall distribution if the error is small (Fig. 5a), to a wide shallow distribution if the error is large (Fig. 5b). These individual distributions are summed together to form the age probability density distribution function, *f(t)*, for the sample being examined (e.g. Fig. 5c), using the following formula:

$$f(t) = \sum_{i=1}^{N} \frac{1}{e_i \sqrt{2\pi}} \exp^{-(t-x_i)^2/2e_i^2} \qquad (11)$$

where x_i is the i^{th} age measurement and e_i is the i^{th} analytical error, t is the age and N is the sample size. In practice, the distribution function can be approximated by assessing the value of *f(t)* in fixed increments (typically 1 Ma) across a range that completely encompasses the required data. Because no simple assumption about the underlying sample distribution can be made, the use of 1 Ma increments can only be tested against the function defined by smaller increments (e.g. 0.1 Ma) by assessing the deviation of interpolated values.

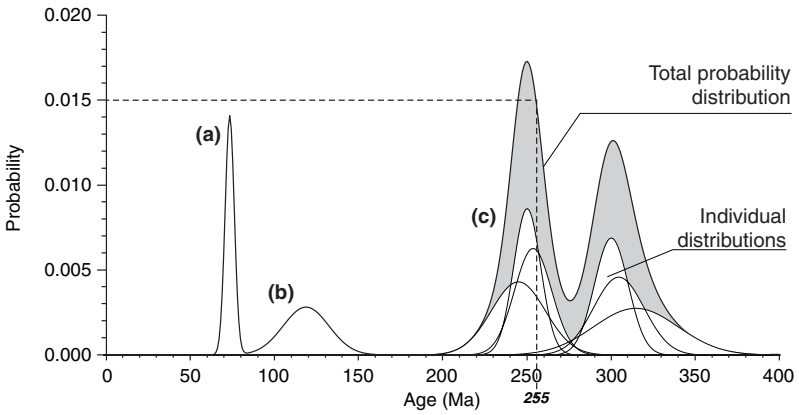

Figure 5.

Example of individual and accumulated ageestimates. ***a****) Individual age estimate with a small error;* ***b****) age estimate with a large error;* ***c****) accumulated density distribution with six contributing age estimates in two modes. Although it is the area beneath the curve that is important (see text) the probability distribution can be read as follows: the probability of an age of 255 Ma within the distribution is 0.015 or 1.5%.*

Using the Dwyer Lake sample data, it was found that deviation in the function using 0.1 Ma and 1 Ma increments was trivial ($R^2 = 0.9993$). For comparison, the function defined by a 10 Ma increment has a correlation of 0.9413, and 25 Ma increments yield a correlation of 0.6715. Because of this and the fact that most SHRIMP geochronological results are usually reported to a round Ma value, standard increments of 1 Ma are considered a suitable approximation for the distribution function. In some cases, particularly involving younger ages, the size of these increments may have to be reduced further to ensure a smooth appearance. It is important to note that for the distribution to be a true probability distribution, it must be scaled so that the cumulative total sums to one, as follows:

$$f(t) = \frac{1}{N} \int_0^{\infty} \sum_{i=1}^{N} \frac{1}{e_i \sqrt{2\pi}} \exp^{-(t-x_i)^2 / 2e_i^2} \, dt = 1 \quad (12)$$

The appropriate scaling in equation (11) can be also achieved by ratioing by N and the cumulative probability function can be approximated by summing each of the 1 Ma steps. This also ensures that the diagram is standardized for comparative purposes. It is also good practice to retain the probability scale on the y-axis to allow meaningful comparison between sets of data. The number of individual analyses contributing to the total distribution should also be clearly indicated on the diagram.

The age probability density distribution diagram counters the two limitations of the binned frequency histogram. Firstly, the individual age errors are used in calculating the probability density distribution estimation and an overall visual impression of the relative precision of the set of data can be made. Nevertheless, it remains good practice to report the mean or median of the errors in the set of data being displayed. Secondly, by estimating the sample distribution in 1 Ma steps, the potential for altering the appearance of the diagram as bin width varies is eliminated and the diagram is in effect standardized to allow for comparison between sets of age data.

Probability density distribution diagrams are particularly good at displaying the modality of a set of age data, as revealed by the peaks of the distribution curve (e.g. Fig. 6). Nevertheless, care must be taken in drawing geological interpretations of the peaks, in particular in picking out peaks as an exact representation of protosource ages in a set of detrital age data. Overlapping modes may slightly skew the location of neighboring peaks, and a mode with few data may be visually indistinct near a relatively larger mode (the same limitations of visual interpretation also apply to binned frequency histograms). The creation of entirely false peaks and modes by the overlap of near, but significantly different, ages is unlikely because for typical SHRIMP detrital ages (e.g. with 10–15 Ma errors at 1 s.e.) to be considered statistically significantly different, a difference between the ages in the order of 25 to 35 Ma is needed, based on a z-test hypothesis at 95% confidence that the real difference between the two ages is not zero. Such a difference is typically seen in a probability density distribution as either distinctive peaks or a pronounced shoulder on a larger peak.

Defining and assessing the significance of modes within multimodal density distributions is a sophisticated process (e.g. Good and Gaskins, 1980; Izenman and Sommer, 1988) that is largely beyond the scope of this paper and will only be discussed briefly. For the purposes of initial data analysis, modes in an age probability density distribution can be defined in terms of the function derivatives as those points

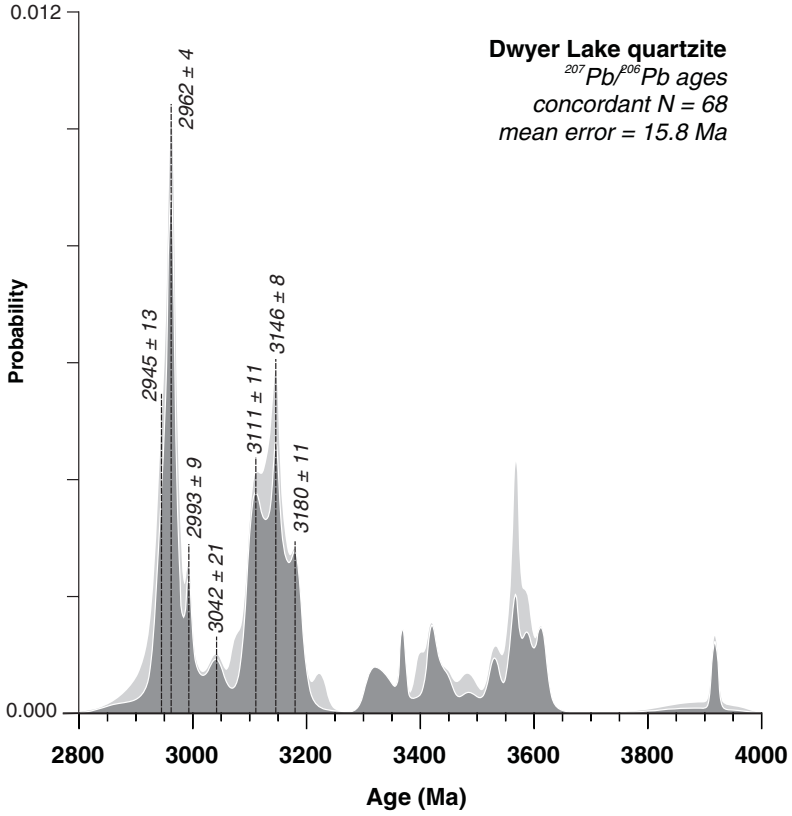

Figure 6.

Dwyer Lake quartzite age data represented using a probability density distribution diagram. Lighter shading indicates distribution with all ages (concordant plus discordant, N=87), darker shading indicates distribution of ages considered concordant (95–105%, N= 68). Vertical dashed lines and ages indicate mixture model ages given in Table 3.

Table 3. Components within the Dwyer Lake quartzite data as indicated by the mixture modelling deconvolution method of Sambridge and Compston (1994). The wide range of ages older than 3200 Ma deconvolve to individual ages and are not listed here.

Component age (Ma)	± error (1 s.e.)	ages in component	relative proportion
2945	13	3	4.4%
2962	4	15	22.0%
2993	9	2	2.9%
3042	21	3	4.4%
3111	11	8	11.8%
3146	8	10	14.7%
3180	11	5	7.4%

where $f'(t)=0$ and $f''(t)<0$ (Scott, 1992). For objectively deriving modal ages from a set of age data, the analyst should use a deconvolution methodology, such as the maximum-likelihood mixture modelling of Sambridge and Compston (1994). Starting with an estimate of the number of components within the set of age data, the mixture modelling algorithm attempts to model the sample distribution by varying the proportion and variance of the modelled components until the best statistical fit is found. The process is continued with increasing numbers of assumed components until the increase in goodness of fit is trivial or, more typically, components begin to repeat. Because of its mathematical basis, this approach avoids any potential bias in visual inspection, although it is interesting to note, as in the case of the Dwyer Lake sample (Table 3, Fig. 6), these calculations often define components that closely match the prominent peaks seen in the probability density distribution. The geological interpretation of these model components and peaks should be weighted carefully in terms of the number of individual measurements making up each component. As a general rule of thumb any modelled component should consist of three or more individual measurements before being considered meaningful. In the Dwyer Lake data, the modelled components at 2962 ± 4 Ma (n=15) and 3146 ± 8 Ma (n=10) should be given greater significance in geological interpretation than the 2993 ± 9 Ma (n=2) component. At this point, data assessment may pass from the objective rigour of mathematical modelling and statistical significance to more subjective geological interpretation. For example, the Dwyer Lake data contain two >3800 Ma ages. Mathematically these could be dismissed as insignificant, but geologically they are potentially very interesting. Similar significance is often also placed on the youngest analysis in a set of detrital data as an indicator of the maximum age of deposition.

An important limitation of age probability density distribution diagrams is the loss of easily accessible frequency information that can be seen in a traditional binned frequency histograms and is often the instinctive demand of someone viewing such a diagram: they simply want to know, even roughly, how many analyses are in a particular interval. In a binned frequency histogram estimate, this information is conveyed by the height of the columns, but this paradigm does not necessarily follow for probability density distribution diagrams. Because probability density distribution diagrams illustrate probability density, it is the *area beneath the curve* that conveys frequency and proportion information. The height of the curve in a probability density distribution diagram is a function of both quantity and precision rather than simple quantity, i.e. a singular and particularly precise analysis will have a tall peak (e.g. Fig. 5a) that may compare, height-wise, with a peak of accumulated, less precise analyses (e.g. Fig. 5c). For this reason it is important to retain the quantity of individual analyses displayed in a probability density distribution diagram to permit a reasonable comparison between different sets of data. Unfortunately, area is not an intuitively recognized attribute of a diagram and frequency information may be lost to a casual audience. For example, in Figure 5c the distribution contains six individual analyses, three in each mode, but the left-hand peak is higher, conveying on first glance that it is the dominant mode, whereas in reality it has exactly the same quantity as the other mode. The height difference is a result of the three analyses on the left being more precise.

There are numerous cases where the display of frequency information is not required, for example, if the intention is to simply display a series of data for comparison where the presence or absence of particular modes is of primary interest (e.g. Geslin et al., 1999). If frequency information is required, overlaying the probability density distribution diagram with an equivalent binned frequency histogram of the same set of age data is recommended. The binned frequency histogram (with bin widths chosen and justified as explained above) will convey frequency information whereas the probability density distribution will constrain the histogram by illustrating the overall precision of the analyses and by providing an unbiased distribution for interpretation. Scaling the probability density distribution to match the area represented by the equivalent histogram simply involves multiplying the probability distribution equation (11) by the number of samples N and histogram bin width h (Bevington and Robinson, 1992):

$$f_{scaled}(t) = f(t)Nh \tag{13}$$

In practice, clarity of presentation may result in modifying the scaling for one or the other of the overlying diagrams. In all cases it is recommended that the sample size, mean/median error, and efficiency of the selected bin width also be displayed or included in the diagram description.

CONCLUSIONS

Absolute age data are displayed to convey the salient features of the sample data (e.g. modes, ranges, and proportions) for both data analysis and communication of interpretations where concordia diagrams may be unsuitable (Table 4). In mathematical terms, displays of univariate age data such as binned frequency histograms and probability density diagrams are estimates of the distribution of the sample. They are often very effective in conveying information, although the analyst should be aware that both approaches involve

Table 4. Summary of the advantages and disadvantages of concordia diagrams, binned frequency histograms, and probability density distributions for displaying relatively large sets of age data.

Style of diagram	Advantages	Disadvantages
Concordia diagrams	• Display of precision and accuracy of results, allows for visual assessment of concordance, Pb-loss. • Display of analytical errors. • Widely used among specialists.	• Quickly becomes visually cluttered as sample size increases. • Difficult to assess for modality and proportions. • Potentially difficult for nonspecialist or nontechnical audience.
Binned frequency histograms	• Effective communication of salient features of data: modes, ranges, and proportions. • Widely used by both specialist and nonspecialist audiences.	• Appearance and interpretation vulnerable to bias because bin width is arbitrary. Not necessarily a widely appreciated issue. • Analytical errors discarded, places all emphasis on calculated age.
Probability density distribution	• Effective communication of modality, range, and overall precision of data. • Uses analytical errors in data and is thus mathematically a better estimate of sample distribution. • Standardization of appearance.	• Potential confusion and misinterpretation because area equals frequency rather than height alone. • Potentially difficult for nontechnical audiences.

limitations and potential for misinterpretation, especially if the diagram is the principal means of communication. In particular, the following points are considered important:

– The analyst should carefully consider the purpose of and the audience for the display. Decisions about the type of display and information conveyed should be made on the basis of questions such as: Is salient sample information about modes, ranges, and proportion sufficient? Will the audience demand and/or have access to analytical information about accuracy and precision?

– In the transition from a bivariate concordia plot to a univariate age data display, visual information regarding the accuracy of the measurements will be lost. Therefore the first step in constructing a display of age data from U-Pb isotopic measurements is to filter the data for accuracy. For detrital SHRIMP analyses, this is typically an arbitrary constraint on concordance such as 95% to 105%. In some cases, such as material less than ~1500 Ma, this filtering has to be judged carefully. The method of filtering should be stated in the diagram description and in some cases it may be useful to display both concordant and discordant data.

– The most widely used estimate is that of the binned frequency histogram, but this display has two major limitations when applied to absolute age data. Firstly, analytical errors are discarded so in effect the bins only count the age. Therefore visual information conveying analytical precision is lost. Diagram description should include an assessment of the mean or median analytical error of the sample.

– The second limitation of binned frequency histograms involves the arbitrary selection of bin width. The appearance, and thus potential interpretation, of the histogram itself is strongly controlled by the bin width parameter and, to a lesser extent in age data display, by the origin. A variety of methods may be applied to find an optimal bin width for the data (e.g. equations (4) and (8); Doane, 1985; Scott, 1979, 1992) but these assume a sample distribution that may not be valid. An alternative approach is proposed (equation 9) where the efficiency of a particular bin width at representing the individual age estimates is calculated. For typical thermal ionization mass spectrometer data with analytical errors of about 1 Ma (1 s.e.), bin widths of 10 Ma will reach 90% efficiency (Table 2). For typical SHRIMP analytical errors at 10 to 15 Ma (1 s.e.), bin widths of at least 20 Ma are needed to reach over 50% efficiency and of much larger than 100 Ma to reach 90% efficiency (Table 2). The diagram description should include a justification of the chosen bin width and a discussion of how the appearance of the distribution varies with other bin widths.

– Probability density distribution diagrams counter the two limitations of binned frequency histograms. A standardization of display can also be achieved by assuming that 1 Ma increments are a suitable approximation of the sample distribution, although in some cases the increment may be smaller. It is good practice to include sample size and mean/median sample error in the diagram description.

– An important limitation of probability density distributions is the loss of visual frequency information. To convey the maximum amount of detailed information about a sample, it may be beneficial to combine the equivalent probability density distribution and binned frequency histogram.

– Care must be taken in the geological interpretation of modes displayed in both binned frequency histograms and probability density distributions. Overlapping modes may slightly skew peaks and numerically smaller modes may be visually indistinct. If detailed derivation of sample modes are required, the analyst should use a mixture modelling method.

ACKNOWLEDGMENTS

This manuscript was greatly improved by the reviews and discussions of B. Davis, R. Stern, and many others. The author is funded by a Canadian Government Laboratory Visiting Fellowship.

REFERENCES

Bevington, P.R. and Robinson, D.K.
1992: Data Reduction and Error Analysis for the Physical Sciences (second edition); McGraw-Hill Ltd., New York, 328 p.

Bleeker, W., Stern, R., and Sircombe, K.
2000: Why the Slave Province, Northwest Territories, got a little bigger; *in* Current Research 2000-C; Geological Survey of Canada, Current Research-C2, 9 p. (online; http://www.nrcan.gc.ca/gsc/bookstore)

Davis, D.W., Hirdes, W., Schaltegger, U., and Nunoo, E.A.
1994: U-Pb age constraints on deposition and provenance of Birimian and gold-bearing Tarkwaian sediments in Ghana, West Africa; Precambrian Research, v. 67, p. 89–107.

Doane, D.P.
1985: Aesthetic frequency classifications; The American Statistician, v. 30, p. 181-183.

Dodson, M.H., Compston, W., Williams, I.S., and Wilson, J.F.
1988: A search for ancient detrital zircons in Zimbabwean sediments; Geological Society of London, Journal, v. 145, p. 977–983.

Full, W.E., Ehrlich, R., and Kennedy, S.K.
1984: Optimal configuration and information content of sets of frequency distributions; Journal of Sedimentary Petrology, v. 54, p. 117–126.

Galbraith, R.F.
1989: The radial plot: graphical assessment of spread in ages; Nuclear Tracks and Radiation Measurement, v. 17, p. 207-214.

Gehrels, G.E. and Dickinson, W.R.
1995: Detrital zircon provenance of Cambrian to Triassic miogeoclinal and eugeoclinal strata in Nevada; American Journal of Science, v. 295, p. 18–48.

Geslin, J.K, Link, P.K, and Fanning, C.M.
1999: High-precision provenance determination using detrital-zircon ages and petrography of Quaternary sands on the eastern Snake River Plain, Idaho; Geology, v. 27, p. 295–298.

Good, I.J. and Gaskins, R.A.
1980: Density estimation and bump-hunting by the penalized likelihood method exemplified by scattering and meteorite data; Journal of the American Statistical Association, v. 75, p. 42–56.

Harley, S.L. and Black, L.P.
1997: A revised Archaean chronology for the Napier Complex, Enderby Land, from SHRIMP ion-microprobe studies; Antarctic Science, v. 9, p. 74–91.

Izenman, A.J. and Sommer, C.J.
1988: Philatelic mixtures and multimodal densities; Journal of the American Statistical Association, v. 83, p. 941–953.

Morton, A.C., Claoué-Long, J.C., and Berge, C.
1996: SHRIMP constraints on sediment provenance and transport history in the Mesozoic Statfjord Formation, North Sea; Geological Society of London, Journal, v. 153, p. 915–929.

Pell, S.D., Williams, I.S., and Chivas, A.R.
1997: The use of protolith zircon-age fingerprints in determining the protosource areas for some Australian dunes sands; Sedimentary Geology, v. 109, p. 233–260.

Rainbird, R.H., Stern, R.A., Khudoley, A.K., Kropachev, A.P., Heaman, L.M., and Sukhorukov, V.I.
1998: U-Pb geochronology of Riphean sandstone and gabbro from Southeast Siberia and its bearing on the Laurentia-Siberia connection; Earth and Planetary Science Letters, v. 164, p. 409–420.

Roback, R.C. and Walker, N.W.
1995: Provenance, detrital zircon U-Pb geochronology, and tectonic significance of Permian to Lower Triassic sandstone in southeastern Quesnellia, British Columbia and Washington; Geological Society of America, Bulletin, v. 107, p. 665–675.

Ross, G.M., Parrish, R.R., and Dudás, F.Ö.
1991: Provenance of the Bonner Formation (Belt Supergroup), Montana: insights from U-Pb and Sm-Nd analyses of detrital minerals; Geology, v. 19, p. 340–343.

Ross, G.M., Parrish, R.R., and Winston, D.
1992: Provenance and U-Pb geochronology of the Mesoproterozoic Belt Supergroup (northwestern United States): implications for age of deposition and pre-Panthalassa plate reconstructions; Earth and Planetary Science Letters, v. 113, p. 57–76.

Sambridge, M.S. and Compston, W.
1994: Mixture modelling of multi-component data sets with application to ion-probe zircon ages; Earth and Planetary Science Letters, v. 128, p. 373–390.

Scott, D.W.
1979: On optimal and data-based histograms; Biometrika, v. 66, p. 605–610.
1992: Multivarite Density Estimation: Theory, Practice and Visualization; Wiley, New York, 376 p.

Scott, D.J. and Gauthier, G.
1996: Comparison of TIMS (U-Pb) and laser ablation microprobe ICP-MS (Pb) techniques for age determination of detrital zircons from Paleoproterozoic metasedimentary rocks from northeastern Laurentia, Canada, with tectonic implications; Chemical Geology, v. 131, p.127–142.

Silverman, B.W.
1986: Density Estimation for Statistics and Data Analysis; Chapman and Hall, London, England, 175 p.

Simonoff, J.S. and Udina, F.
1997: Measuring the stability of histogram appearance when the anchor position is changed; Computational Statistics and Data Analysis, v. 23, p. 335–353.

Sircombe, K.N.
1999: Tracing provenance through the isotope ages of littoral and sedimentary detrital zircon, eastern Australia; Sedimentary Geology, v. 124, p. 47–67.

Sircombe, K.N. and Freeman, M.
1999: Provenance of detrital zircon on the Western Australian coastline — implications for the geological history of the Perth Basin and denudation of the Yilgarn Craton; Geology, v. 27, p. 879–882.

Sturges, H.A.
1926: The choice of a class interval; Journal of the American Statistical Association, v. 21, p. 65–66.

Tera, F. and Wasserburg, G.J.
1972: U-Th-Pb systematics in three Apollo 14 basalts and the problem of initial Pb in lunar rock; Earth and Planetary Science Letters, v. 14, p. 281–304.
1974: U-Th-Pb systematics in lunar rocks and inferences about lunar evolution and the age of the moon; *in* Proceedings, Fifth Lunar Science Conference, v. 2, p. 1571–1599.

Torley, R.K.
1998: Fourier grain-shape analysis of detrital quartz grains in sediments of Bear Creek valley and adjacent western Cascade Mounatins, southern Oregon; Ph.D. thesis, University of Oregon, Oregon, 264 p.

Tukey, J.W.
1977: Explanatory Data Analysis; Addison-Wesley, Reading, 688 p.

Wand, M.P.
1996: Data-based choice of histogram bin width; University of New South Wales, Australian Graduate School of Management, Working Paper Series 95-011, 14 p.

Wetherill, G.W.
1963: Discordant uranium-lead ages, Part 2, discordant ages resulting from diffusion of lead and uranium; Journal of Geophysical Research, v. 68, p. 2957–2965.

Whitehouse, M.J., Claesson, S., Sunde, T., and Vestin, J.
1997: Ion microprobe U-Pb zircon geochronology and correlation of Archean gneisses from Lewisian Complex of Gruinard Bay, northwestern Scotland; Geochimica et Cosmochimica Acta, v. 61, p. 4429–4438.