[AMERICAN JOURNAL OF SCIENCE, VOL. 292, OCTOBER, 1992, P. 535–564]

American Journal of Science

OCTOBER 1992

DECOMPOSITION OF FISSION-TRACK GRAIN-AGE DISTRIBUTIONS

MARK T. BRANDON

Department of Geology and Geophysics, Yale University,
P.O. Box 6666, New Haven, Connecticut 06511

ABSTRACT. The external-detector method of fission-track (FT) dating is capable of providing FT ages for individual mineral grains. As a result, this method when applied to dating detrital zircons in unreset sandstone samples has much potential for resolving the depositional age and provenance of the samples. One problem is that individual FT grain ages for detrital zircons usually have fairly low precision, with an average relative standard error of about 13 percent. As a result, it is necessary to find related groups of grain ages in order to improve the precision of the age estimates.

This paper introduces two new methods for decomposing a FT grain-age distribution from a single sandstone sample into component grain-age populations. The FT grain ages from a sample of unreset detrital zircons reflect the cooling and/or magmatic history of the source region from which they were eroded. The sample grain-age distribution can be viewed as a mixture of component populations, each of which is related to a FT source terrain with a characteristic FT age. The first decomposition method, the χ^2 age method, isolates the youngest fraction of "plausibly related" grain ages and assigns an age to this fraction, called the χ^2 age. This age estimate is useful in that it provides a maximum limit for the depositional age of the sandstone sample. The second method, the Gaussian peak-fitting method, decomposes the entire grain-age distribution into a finite set of component Gaussian distributions, each of which is defined by a unique mean age, a relative standard deviation, and an estimated number of grain ages in the component distribution. A series of simulation experiments indicates that both these methods will produce satisfactory results given a sufficient number of grain ages for each peak and adequate separation between adjacent peaks. A practical application of these methods is given in a companion paper (Brandon and Vance, this issue).

INTRODUCTION

When using the external-detector method of fission-track (FT) dating (Naeser, Naeser, and McCulloh, 1989), a FT chronologist actually determines the ages of individual mineral grains. Usually about 10 to 20 grains are dated per sample, and their ages are grouped together to produce a more precise average age for the sample. The assumption here is that the dated grains have a common thermal history, as might be

expected for a group of zircon grains in an unmetamorphosed volcanic tuff.

Another less common application of the FT method is to date detrital grains from sedimentary rocks, most commonly zircon from sandstones (Hurford, Fitch, and Clarke, 1984; Baldwin, Harrison, and Burke, 1986; Kowallis, Heaton, and Bringhurst, 1986; Cerveny and others, 1988; Naeser, Naeser, and McCulloh, 1987, 1989; Brandon and Vance, this issue). The zircon grains in a sandstone sample will retain detrital FT ages as long as the sandstone has remained at fairly low temperatures after deposition, less than about 170° to 180°C, the estimated threshold for the onset of significant annealing of zircon fission tracks (see Brandon and Vance, this issue). Thus, these detrital FT ages should reflect the cooling history and/or magmatic history of the source region from which the dated mineral grains were eroded.

Obviously, single grain dating methods have great potential for providing new information about the provenance and age of unreset sedimentary rocks. The FT system is particularly well suited for this application because of its low closure temperature. The closure temperature for fission tracks in zircon is estimated to be about 235° to 245°C assuming typical cooling rates of 10° to 30°C/my (see Brandon and Vance, this issue). Thus, FT ages generally reflect the thermal history of rocks as they cool within the upper part of the Earth's crust. This is important because, relative to other higher temperature isotopic systems (K-Ar for hornblende or white mica, U/Pb for zircon), the amount of time between closure of the FT system in rocks of the source region and deposition of sediment derived from those rocks in a sedimentary basin should be relatively short. This time lag might be further shortened by tectonic and/or magmatic activity in the source region.

This work with FT detrital grain ages has been driven by two objectives. The first was to see if the method could be used to resolve better the depositional age of poorly fossiliferous sedimentary units in the Cenozoic Olympic subduction complex of Washington State (Brandon and Vance, this issue). The governing principle here is that the depositional age of a rock can be no older than the age of its detrital components. Therefore, the age of the youngest fraction of FT grain ages represents the maximum depositional age of an unreset sample. The second objective was to characterize better the components or peaks in a grain-age distribution in order to derive information about the provenance of the sample. It also turns out that the older peaks in a detrital grain-age distribution can sometimes be use to determine if a sample has been partially reset because this process will reduce the apparent age of the peaks (Brandon and Vance, this issue).

A significant impediment to this kind of study is the analysis of the resulting grain-age distributions. The method tends to generate a large amount of data, with a typical distribution for a single sample containing 50 or more grain ages. The more serious problem, however, is that an individual grain age is generally not very precise. The average uncer-

tainty for a zircon FT grain age is about ±26 percent of the age at ±2s (two standard error, equal to a confidence limit of about 95 percent). Thus, it becomes essential to gather grain ages into related groups in order to improve the precision of our age determinations. When a group of grain ages can be averaged, the standard error is decreased by a factor of $1/\sqrt{N}$, where N equals the number of grains. For example, for a group of 10 related grains, ±2s would be about ±8 percent of their average age. This level of uncertainty is quite reasonable for rocks with young grain-age distributions, less than about 50 Ma.

One approach to this problem is to view a grain-age distribution as a mixture of component populations, with each component derived from a specific FT source terrain. The term *FT source terrain* is used here to refer to a discrete area in the general source region that yields zircons of a characteristic FT age. (Note that this term uses the American spelling for the word "terrain" to avoid confusion with the term "tectonic terrane".) In reality, a given source region is probably capable of delivering a wide range of FT grain ages, but it is anticipated that a few dominant component populations will account for the bulk of the grain ages.

In statistics, this problem falls under the general topic of finite-mixture distributions. There are several recently published books (Everitt and Hand, 1981; Titterington, Smith, and Makov, 1985; and McLachlan and Basford, 1988) that discuss the properties of finite-mixture distributions and methods for decomposing them into component populations. The more specific problem of decomposing mixed distributions of FT grain ages has been addressed by Seward and Rhoades (1986), who propose a method based on cluster analysis, and Galbraith (1988) and Galbraith and Green (1990), who introduce some decomposition methods based on the maximum likelihood method.

Two new methods are introduced here. The first, the χ^2 age method, is designed to isolate the youngest fraction of "plausibly related" grain ages in a grain-age distribution from a single sandstone sample. The age of this fraction, which is called the χ^2 age, is useful in that it provides a maximum limit for the depositional age of an unreset sample.

The second method, the Gaussian peak-fitting method, actually decomposes the grain-age distribution into component Gaussian distributions. This method has some general similarities with that of Galbraith and Green (1990), but the implementation is quite different. In a comparison given below, it is shown that these two peak-fitting methods produce fairly similar results. It is acknowledged that the method of Galbraith and Green (1990) is statistically quite elegant. However, the Gaussian peak-fitting method has several practical advantages: (1) The method makes fewer assumptions about the structure of the data and allows a greater degree of choice in selecting model parameters. (2) The numerical algorithm may be better suited for finding multiple peaks, especially when 4 or more peaks are present. (3) The method of solution is more

intuitive and provides some useful insights into the problems associated with decomposing FT grain-age distributions.

This paper assumes that the reader has a general knowledge of the procedures and statistical methods used in FT dating. Useful background references are Gleadow (1981), Green (1981), Naeser, Naeser, and McCulloh (1989), and Hurford (1990).

χ^2 AGE METHOD

An important feature of the FT method is that it is possible to estimate the uncertainty associated with each age determination. The reason is that the radioactive decay process follows a Poisson distribution with the number of counted tracks being the direct manifestation of the decay process. As a result, FT dating of a single sample produces a set of estimated grain ages μ_i and associated standard errors s_i for a total of N_t grains with $i = 1, \ldots, N_t$. If the sample being dated is considered to be composed of a single population of FT grain ages, such as zircon from a volcanic tuff, then we would be interested in calculating an average for the grain ages. According to accepted procedure, the χ^2 test of Galbraith (1981) and Green (1981) is used to determine how best to estimate this average. The χ^2 test compares the observed variance within the population of grain ages with the variance predicted for the radioactive decay process alone. The test returns $P(\chi^2)$, the probability that the difference between observed and predicted variances could be due to random chance alone. The test is considered to be passed when $P(\chi^2) > 5$ percent. In this case, the age of the grains is represented by a *pooled age*, which is calculated by pooling the track count data for all grains. This procedure is akin to calculating a weighted mean. If the χ^2 test is failed, then a *mean age* is used, which is calculated as the unweighted mean of the grain ages. Note that neither of these methods is suitable for calculating an age for a mixed distribution of FT grain ages. However, the χ^2 test can be used to isolate the youngest fraction of "plausibly related" grain ages.

The first step in the χ^2 age method is to sort the grains in order of increasing age. An inspection of the example shown in table 1 reveals that the first 10 to 15 grains have similar grain ages, and the remaining 35 to 40 grains appear to be much older. The dividing point, however, is difficult to resolve because the estimated uncertainties for the grain ages are large. The objective is to isolate a group of young grain ages that with statistical confidence could conceivably have been derived from a single-age source.

A modified application of the χ^2 test of Galbraith (1981) and Green (1981) is used for this purpose. With each grain age, we calculate a sum age and the χ^2 probability, $P(\chi^2)$. The sum age is the pooled age for all grain ages equal to or younger than the current grain age. $P(\chi^2)$ indicates the probability that the variability in age within the current fraction of young grain ages is compatible with that expected for grain ages from a single-age source. $P(\chi^2)$ is calculated using the formula given in Green (1981, p. 81) and an algorithm given in Press and others (1986).

TABLE 1

Grain ages for Mount Tom (ZD6)

No.	Grain age ± 2s (Ma)		P(χ^2) (%)	Sum age ± 2s (Ma)	
1	14.0	4.6	100.0	14.0	4.6
2	14.1	4.1	96.6	14.0	3.1
3	14.1	7.3	99.9	14.1	2.8
4	17.4	6.3	78.5	14.8	2.6
5	18.0	6.0	69.7	15.4	2.4
6	18.9	6.8	65.0	15.9	2.3
7	19.2	8.3	67.4	16.2	2.3
8	20.0	6.2	58.8	16.7	2.2
9	20.0	6.9	58.6	17.1	2.1
10	20.1	9.2	63.5	17.2	2.0
11	20.6	6.5	61.3	17.6	2.0
12	20.9	7.2	61.7	17.8	1.9
13	21.3	6.3	58.6	18.2	1.9
14	28.2	9.7	22.3	18.8	1.9
15	**31.8**	**16.3**	**11.1**	**19.1**	**1.9***
16	35.8	12.0	0.3	19.9	1.9
17	38.7	12.6	0.0	20.8	1.9
18	38.9	9.6	0.0	22.2	1.9
19	40.3	14.1	0.0	22.9	2.0
20	40.8	10.9	0.0	23.9	2.0
21	41.0	12.9	0.0	24.6	2.0
22	41.5	11.8	0.0	25.4	2.0
23	42.8	12.0	0.0	26.2	2.0
24	44.5	12.4	0.0	27.0	2.0
25	44.8	15.0	0.0	27.5	2.0
26	45.7	13.5	0.0	28.2	2.1
27	45.8	11.7	0.0	29.0	2.1
28	46.3	14.3	0.0	29.5	2.1
29	49.1	10.5	0.0	30.7	2.1
30	50.5	18.6	0.0	31.0	2.1
31	50.6	15.0	0.0	31.6	2.1
32	54.6	18.7	0.0	32.0	2.1
33	55.2	21.9	0.0	32.4	2.1
34	58.7	17.7	0.0	33.0	2.2
35	58.8	14.1	0.0	34.0	2.2
36	59.8	23.7	0.0	34.3	2.2
37	60.7	16.1	0.0	35.1	2.2
38	63.0	20.9	0.0	35.6	2.2
39	63.6	17.5	0.0	36.3	2.2
40	65.3	18.2	0.0	37.0	2.3
41	66.7	18.2	0.0	37.7	2.3
42	72.0	22.4	0.0	38.3	2.3
43	73.0	22.6	0.0	38.9	2.3
44	77.8	29.6	0.0	39.4	2.3
45	80.2	24.5	0.0	40.1	2.4
46	81.2	28.3	0.0	40.6	2.4
47	81.4	29.0	0.0	41.1	2.4
48	91.8	30.8	0.0	41.8	2.4
49	153.1	55.4	0.0	42.9	2.5
50	195.0	62.9	0.0	44.7	2.6
POOLED AGE				44.7	2.6
MEAN AGE				50.2	9.7

* χ^2 age = 19.1 ± 1.9 Ma.

The χ^2 age is defined here as the pooled age of the largest group of young grains that still retains $P(\chi^2)$ greater than 1 percent. For the sample shown in table 1, the χ^2 age is 19.1 Ma and corresponds to the sum age for the first 15 grains. The reasoning here is that a group of young grains with $P(\chi^2)$ less than 1 percent would probably not have been derived from a single-age source. Experience indicates that the sum age usually changes very slowly when $P(\chi^2)$ is between 5 to 1 percent. Thus, for most cases, the choice of a critical probability has very little effect on the χ^2 age. Nonetheless, there are several reasons to favor a conservative critical probability: (1) The χ^2 test is repeated many times. A smaller critical probability reduces the number of tests that fail due to chance alone (a false negative result). (2) A subtle bias is introduced by ranking grain ages with increasing age. Anomalously young ages in the youngest group of related grains might cause the χ^2 age to underestimate the pooled age of that group. A smaller critical probability helps to offset this potential bias. (3) In the strict sense, unreset grain ages provide a maximum limit for the depositional age of the sandstone. Thus, if a bias exists, it is preferable that the χ^2 age slightly overestimates the pooled age of the youngest group of related grains.

GAUSSIAN PEAK-FITTING METHOD

The Gaussian peak-fitting method decomposes the grain-age distribution into a set of component grain-age populations. This method utilizes the composite probability density plot of Hurford, Fitch, and Clarke (1984), which portrays the grain-age distribution as a continuous probability density plot (fig. 1). This plot is described in more detail below as a first step toward introducing the peak-fitting method.

Composite probability density plots.—Once again, we start with a set of estimated grain-ages μ_i and associated standard errors s_i with $i = 1, \ldots, N_t$. To calculate the composite probability density plot for this distribution, each grain age is recast as a Gaussian probability density function,

$$p_i(x) = \frac{1}{\sqrt{2\pi}\, s_i} \exp\left[-\frac{1}{2}\left(\frac{x - \mu_i}{s_i}\right)^2\right] \tag{1}$$

where $p_i(x)$ is the probability density (Ma^{-1}) for the ith grain at time x (Ma). The composite probability density plot is simply the sum of all the grain-age Gaussian functions divided by N_t,

$$p_t(x) = \frac{1}{N_t} \sum_{i=1}^{N_t} p_i(x) \tag{2}$$

where $p_t(x)$ is the composite probability density at time x. Remember that μ_i and s_i are estimated using a Poisson distribution. A Gaussian distribution can be substituted as long as there are more than about 9 spontaneous tracks and 9 induced tracks per grain (Taylor, 1982, p. 120), which is generally the case for FT dating of zircon.

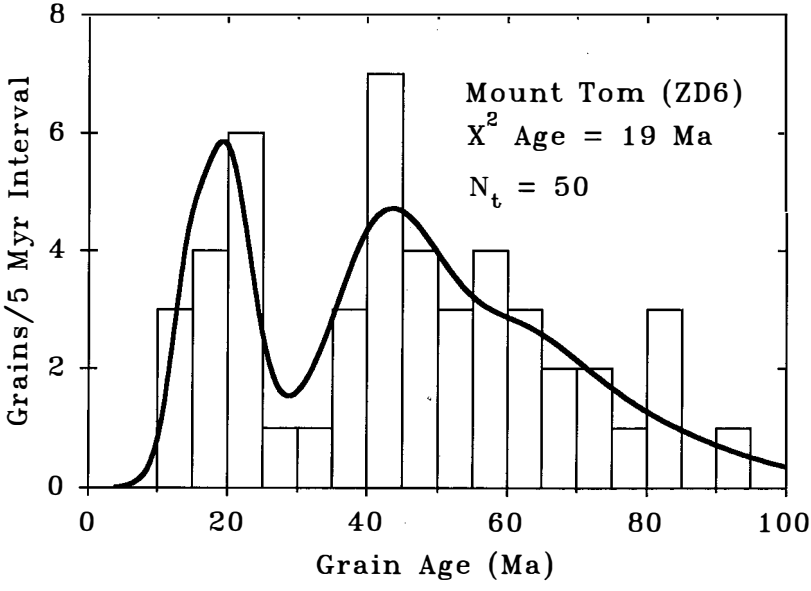

Fig. 1. An example of a composite probability density plot and grain-age histogram for sample ZD6 from table 1.

At this point, it becomes important to define s_i more accurately. There are two types of uncertainty—grain error and group error—that contribute to the total uncertainty assigned to an estimated FT age. *Grain error* corresponds to the uncertainties associated with measurements made on individual grains, that is, the measurement of spontaneous and induced track densities in each grain. *Group error* corresponds to those uncertainties that are common to all grains, such as the uncertainties associated with a laboratory calibration constant, such as the zeta factor, or the measured track density in the mica detector of a fluence monitor. These measured parameters are not expected to vary within a single irradiated sample, but their precision does affect the precision of any estimated FT age. In Brandon and Vance (this issue), the average total relative standard errror for a grain age is ~13.3 percent. Of this total, the relative standard errors due to grain error and group error are ~13.1 and ~2.35 percent, respectively.

The important point here is that the s_i values used in the composite probability density plot (eq 1) should include only grain errors. Note, however, that the estimated uncertainties for grain ages reported in table 1 include both grain and group errors because we are interested in the total uncertainty of those ages. Likewise, the group error must be added to the estimated uncertainty of each peak-fit age (see eq 15), as determined by the Gaussian peak-fitting method, if this estimate is to represent the total uncertainty of the peak age.

An important advantage of the composite probability density plot is that grain ages are represented by a continuous function (eq 2). Furthermore, each grain age is distributed relative to its associated uncertainty. A grain age with a large uncertainty contributes a relatively small probability density at any specific age, because the probability mass is spread over a large age range. Conversely, a grain age with a small uncertainty contributes a larger probability density at any specific age, because the probability mass is concentrated over a smaller age range. These attributes provide a significant advantage over a histogram, the appearance of which is strongly influenced by the size and position of the histogram intervals.

For purposes of comparison, a grain-age histogram is superimposed on the composite probability density plot shown in figure 1. This result was accomplished by scaling the composite probability density $p_t(x)$ by a factor of $5N_t$ so that it has the same units as the frequency variable in the histogram (number of grains per 5 my interval).

Decomposing finite-mixture distributions.—Following McLachlan and Basford (1988, p. 9), a group of FT grain ages from an unreset sandstone sample can be viewed as a superpopulation that contains a mixture of several different grain-age populations. In the ideal case, each grain-age population would correspond to a discrete source terrain with a characteristic FT age. In our application, we are interested in estimating the statistical properties of these component populations using grain age data from a sample grain-age distribution. A separate issue, considered elsewhere, is the relationship of the sample distribution to its parent distribution.

Let's assume that each of the component populations in the composite probability density plot can be represented by a unique Gaussian distribution. In this case, we are interested in determining the mean μ_{fj}, relative standard deviation W_{fj}, and number of grains N_{fj} associated with each of these component populations, where $j = 1, \ldots, g$. Note that W_{fj} is equal to σ_{fj}/μ_{fj}, where σ_{fj} is the absolute standard deviation. The relative standard deviation is preferred here because the absolute standard deviation for a group of FT grain ages from a single-age source increases with increasing age, but the relative standard deviation remains fairly constant (Galbraith, 1990, p. 209). It is important to note that W_{fj} and σ_{fj} measure the relative and absolute standard deviation, respectively, of a component peak in the composite probability density plot. These parameters are only indirectly related to the relative and absolute standard deviations of grain ages in the component peak. This distinction is examined in a subsequent section ("Interpretation of best-fit Gaussian parameters").

In decomposing a composite probability density plot, the trivial solution would be a set of N_t Gaussians, corresponding to the original N_t grain ages used to construct the plot. We are interested in the non-trivial solution where g is much less than N_t.

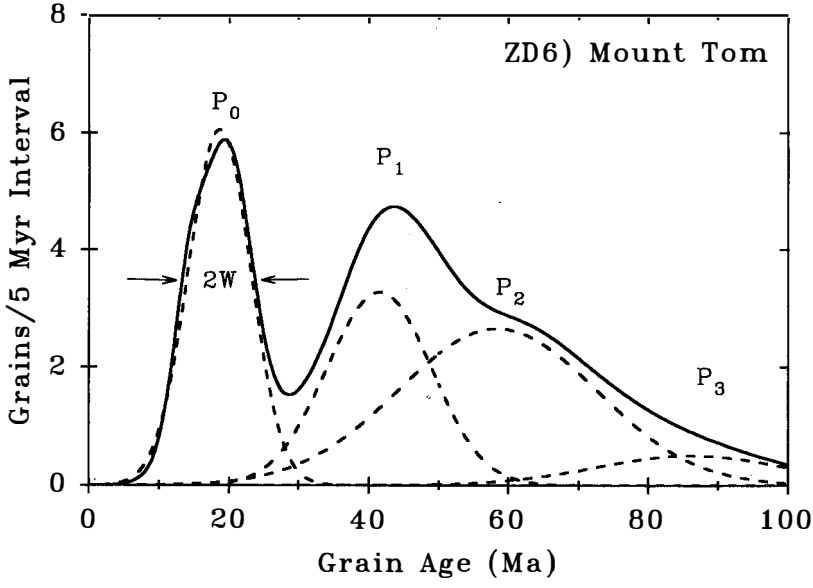

Fig. 2. Best-fit Gaussian peaks for a composite probability density plot (sample ZD6 from table 1). The solid line corresponds to the composite probability density plot, and the dashed lines to the component Gaussian peaks. The peaks are labeled P_0 for the youngest peak and P_1, P_2, and P_3 for the older peaks. The width of a peak is measured by its relative standard deviation, W, as illustrated by the label "2W."

How can we identify the underlying Gaussian distributions? One approach might be to pick prominent peaks in the composite probability density plot (Hurford, Fitch, and Clarke, 1984; Kowallis, Heaton, and Bringhurst, 1986; Cerveny and others, 1988). This method is unreliable because overlap between adjacent Gaussians can cause peaks to shift away from their associated means and can also mask smaller peaks in the distribution (fig. 2); others examples are shown in Everitt and Hand, 1981, p. 27–29; and Titterington, Smith, and Makov and others, 1985, p. 48–50). Note that from here on, the term "peak" is applied to any component Gaussian distribution, even those for which there is no visible peak in the composite probability density plot.

The Gaussian peak-fitting method, described here, fits a series of Gaussian functions to the composite probability density plot using a least-squares criterion. This particular inverse problem is discussed in Menke (1989, p. 210–213), Press and others (1986, p. 528), and Titterington, Smith, and Makov (1985, p. 116, L_2-norm method). The program[1]

[1] On request, the author will provide copies of the source code (Microsoft Professional Basic) for the programs used in this study including an implementation of Galbraith and Green's (1990) binomial peak-fitting method. The copies are available in PC diskette format only; please send a formatted 5.25 inch or 3.5 inch diskette with your request.

that I have developed uses the non-linear Levenberg-Marquardt algorithm given in Press and others (1986, p. 523–528). To implement this algorithm, we need to specify a model function and its derivatives with respect to the parameter variables.

The model function for the Gaussian peak-fitting method is

$$p_t(x)_{calc} = \frac{1}{N_t} \sum_{j=1}^{g} \frac{N_{f_j}}{\sqrt{2\pi}\,\mu_{f_j} W_{f_j}} \exp\left[-\frac{1}{2}\left(\frac{x - \mu_{f_j}}{\mu_{f_j} W_{f_j}}\right)^2\right] \tag{3}$$

which gives the calculated probability density at time x. The objective is to find a set of Gaussian parameters $\{\mu_{f1}, W_{f1}, N_{f1}, \ldots, \mu_{fg}, W_{fg}, N_{fg}\}$ for the g underlying distributions that minimizes the misfit between the observed and calculated composite probability density functions, $p_t(\cdot)_{obs}$ and $p_t(\cdot)_{calc}$ respectively.

In practice, the continuous probability density functions (2,3) are replaced with frequency functions,

$$\phi_t(x)_{obs} = \sum_{i=1}^{N_t} \phi_i(x)_{obs} = \Delta x \sum_{i=1}^{N_t} p_i(x) \tag{4}$$

$$\phi_t(x)_{calc} = \Delta x \sum_{j=1}^{g} \frac{N_{f_j}}{\sqrt{2\pi}\,\mu_{f_j} W_{f_j}} \exp\left[-\frac{1}{2}\left(\frac{x - \mu_{f_j}}{\mu_{f_j} W_{f_j}}\right)^2\right] \tag{5}$$

where $\phi_t(x)_{obs}$ and $\phi_t(x)_{calc}$ represent the observed and calculated number of grains occurring within a specified age interval, Δx, centered on the age x. For the peak-fitting algorithm, these two functions are discretized using a set of ages, $x_k = \{x_1, \ldots, x_m\}$, which are selected at evenly spaced intervals, equal to Δx, over a range slightly larger than the range of observed grain ages. The number of points, m, is set to $3N_t$ given that the function $\phi_t(\cdot)_{obs}$ is based on N_t grain ages with three parameters per grain age: μ_i, s_i, and a probability mass equivalent to one grain.

The Levenberg-Marquardt algorithm uses the partial derivatives of the model function (5) with respect to the fit parameters to search for the best-fit solution. These derivatives are

$$\frac{\partial \phi_t}{\partial \mu_{f_j}} = \Delta x \sum_{k=1}^{m} \frac{N_{f_j}(x_k^2 - \mu_{f_j} x_k - \mu_{f_j}^2 W_{f_j}^2)}{\sqrt{2\pi}\,\mu_{f_j}^4 W_{f_j}^3} \exp\left[-\frac{1}{2}\left(\frac{x_k - \mu_{f_j}}{\mu_{f_j} W_{f_j}}\right)^2\right] \tag{6}$$

$$\frac{\partial \phi_t}{\partial W_{f_j}} = \Delta x \sum_{k=1}^{m} \frac{N_{f_j}(x_k^2 - 2\mu_{f_j} x_k + \mu_{f_j}^2 - \mu_{f_j}^2 W_{f_j}^2)}{\sqrt{2\pi}\,\mu_{f_j}^3 W_{f_j}^4} \exp\left[-\frac{1}{2}\left(\frac{x_k - \mu_{f_j}}{\mu_{f_j} W_{f_j}}\right)^2\right] \tag{7}$$

$$\frac{\partial \phi_t}{\partial N_{f_j}} = \Delta x \sum_{k=1}^{m} \frac{1}{\sqrt{2\pi}\,\mu_{f_j} W_{f_j}} \exp\left[-\frac{1}{2}\left(\frac{x_k - \mu_{f_j}}{\mu_{f_j} W_{f_j}}\right)^2\right] \tag{8}$$

The degree of misfit between the calculated model function and the observed data is measured by a χ^2 merit function (Press and others, 1986, p. 502),

$$\chi^2 = \sum_{k=1}^{m} \left[\frac{\phi_t(x_k)_{obs} - \phi_t(x_k)_{calc}}{s_\phi(x_k)} \right]^2 \tag{9}$$

where $s_\phi(x_k)$ is the standard error for $\phi_t(x_k)_{obs}$. The advantage of this merit function is that the amount of misfit between each observation and the calculated model is scaled relative to the estimated uncertainty for that observation. Thus, the expectation is that χ^2 for a good model—that is a model that fits the observed data well—will be roughly equal to $\nu = m - 3g$, where ν is the number of degrees of freedom for the model. Another advantage is that the χ^2 value can be used to judge the statistical significance of various best-fit solutions, as discussed below.

All this assumes that suitable estimates are available for $s_\phi(x_k)$. Two different methods have been examined. For the first method, it is argued that $\phi_t(x_k)_{obs}$ should follow a Poisson distribution because this variable measures the number of grain ages within a discrete age interval (Bevington, 1969, p. 84–88; Press and others, 1986, p. 470). Based on the Poisson distribution, the best estimate of $s_\phi(x_k)$ should be,

$$s_\phi(x_k) = \sqrt{\phi_t(x_k)_{obs}} \tag{10}$$

The second method assumes that $s_\phi(x_k)$ can be determined by propagating errors through eqs (1) and (4). The assumption made here is that the uncertainty in the estimate of $\phi_t(x_k)_{obs}$ is related to the cumulative uncertainties associated with the estimated FT grain ages. In this case, $s_\phi(x_k)$ can be estimated using the standard formula for propagation of error,

$$s_\phi^2(x_k) = \sum_{i=1}^{N_t} \left(\frac{\partial \phi_i(x_k)}{\partial \mu_i} s_i \right)^2 = \sum_{i=1}^{N_t} \left(\frac{\phi_i(x_k)(x_k - \mu_i)}{s_i} \right)^2 \tag{11}$$

These two methods are based on two very different assumptions. For the Poisson method, the estimated s_ϕ represents the expected variability if we were to go back to the parent distribution and select another sample distribution of grain ages. For the propagation method, the estimated s_ϕ reflects the uncertainties associated with the estimated FT grain ages in the existing sample distribution. As it turns out, both methods produce nearly identical best-fit solutions because the relative weighting of the observed frequency data is similar. However, the standard errors s_ϕ estimated by the two method differ considerably in magnitude, with the Poisson estimates being larger by a factor of 10 or more. As a result, the χ^2 values produced using the Poisson method are smaller by a factor of ~ 0.01 than those produced by the propagation method. In practice, this difference in absolute magnitude is generally not important. However, it does become important if we want to use χ^2 to judge if the degree of misfit between the model and the data is due to unrepresented information or

to statistical noise in the data. Note that the Poisson method was used to estimate s_ϕ in this study and also in Brandon and Vance (this issue).

Initial guess.—The fitting algorithm requires an initial guess for the Gaussian parameters of the g underlying distributions. Toward this end, a useful guide is the composite probability density plot. As illustrated in figure 2, a well separated Gaussian distribution will produce a peak in the plot (P_0), whereas a smaller Gaussian distribution that overlaps significantly with another larger Gaussian distribution is marked only by a subtle shoulder (P_2) (see discussion of modality and bitangentiality in Titterington, Smith, and Makov, 1985, p. 48–50). Both these situations will produce a minimum in the second derivative of the composite probability density function (Titterington, Smith, and Makov, 1985, p. 55), which can be explicitly calculated from (1,2),

$$\frac{d^2p_i(x)}{dx^2} = \frac{(x - \mu_i)^2 - s_i^2}{\sqrt{2\pi}\, s_i^5} \exp\left[-\frac{1}{2}\left(\frac{x - \mu_i}{s_i}\right)^2\right]$$

$$\frac{d^2p_t(x)}{dx^2} = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{d^2p_i(x)}{dx^2} \tag{12}$$

The minima in this function are used to determine an initial guess of the number of underlying distributions and their means μ_{fj}, as illustrated in figure 3. Each of the relative standard deviations W_{fj} is initially set to 16 percent, which represents the average value for W for FT grain-age distributions from single-age volcanic sources (see samples F1 and C5 in Brandon and Vance, this issue). Initial values for N_{fj} are estimated using a modified form of eq (3) where μ_{fj} is substituted for x,

$$N_{f_j} \approx \sqrt{2\pi}\,[p_t(\mu_{f_j})N_t\mu_{f_j}W_{f_j}] \tag{13}$$

Number of peaks.—In most cases, the fitting routine is fairly insensitive to the initial guess used in the search, although the distance between the initial guess and the final solution does affect the amount of time needed to converge on a solution. Usually the most difficult problem is determining g, the number of underlying distributions. The approach adopted here is to search for a terse solution, defined as the minimum number of Gaussians needed to represent fully the composite probability density plot. This approach is implemented by calculating best-fit solutions for an incrementally larger number of Gaussians, starting with one and ending with all the potential Gaussians detected by the second-derivative (see example in table 2). A minimum requirement for a candidate solution is that it must be realistic, that is, μ_{fj}, W_{fj}, and N_{fj} must all be greater than zero, and the peak ages μ_{fj} must be unique. Three other criteria are used: (1) The sum of the estimated number of grains in the component populations, ΣN_{fj}, must approach N_t. (2) The relative standard deviations W_{fj} should approach a value of about 16 percent, which is the expected value for a peak derived from a single-age FT

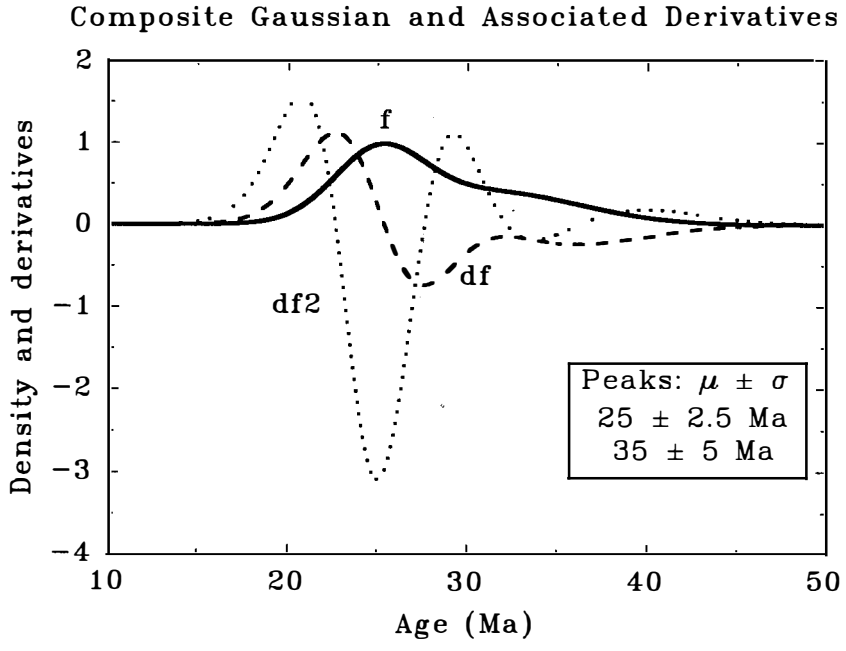

Fig. 3. An example of a composite probability density plot (solid line) and its first and second derivatives (dashed and dotted lines, respectively). Minima in the second derivative are used to make an initial guess of the location of potential peaks in the plot. The example shown here contains two peaks with means and standard deviations of 25 ± 2.5 Ma and 35 ± 5 Ma.

source. (3) The addition of a new peak should result in a significant reduction in the χ^2 value. The level of significance can be determined using the F-ratio test (Bevington, 1969, p. 200). The F-ratio is calculated as

$$F_\chi = \frac{\chi_a^2 - \chi_b^2}{\chi_b^2/\nu_b} \tag{14}$$

where χ_a^2 refers to a model with g-1 peaks, χ_b^2 refers to a subsequent model using g peaks, and ν_b is the number of degrees of freedom for this subsequent model ($\nu_b = 3[N_t - g]$). The objective here is to determine the probability that the incremental change in χ^2 due to the addition of a new peak might be due to random chance alone. This probability $P(F, \nu_1, \nu_2)$ is determined using an F distribution where F is estimated from (14), and the degrees of freedoms for the distribution are $\nu_1 = 1$ and $\nu_2 = \nu_b$. This procedure is illustrated in table 2. Note that in all cases shown, $P(F, 1, \nu) \approx 0$ percent indicating that the introduction of each additional peak resulted in a significant decrease in χ^2.

TABLE 2

An example of an incremental search for the optimal number of Gaussian peaks in a distribution of FT zircon grain ages

Number of peaks	Best-Fit Peaks*					Fit Statistics†	
2	18.1 Ma N_f = 9 W = 20%	-- n.d. --	50.5 Ma N_f = 38 W = 43%	-- n.d. --	-- n.d. --	χ^2 = 3.3 ν = 144 ΣN_f = 47.7	— —
3	18.3 Ma N_f = 11 W = 22%	43.4 Ma N_f = 8 W = 15%	55.4 Ma N_f = 29 W = 40%	-- n.d. --	-- n.d. --	χ^2 = 2.00 ν = 141 ΣN_f = 48.5	P(F,1,ν) = 0% F = 90.3
4	18.4 Ma N_f = 12 W = 23%	42.9 Ma N_f = 8 W = 15%	56.3 Ma N_f = 28 W = 36%	-- n.d. --	169 Ma N_f = 2 W = 25%	χ^2 = 0.13 ν = 138 ΣN_f = 50.2	P(F,1,ν) = 0% F = 1910
5	18.5 Ma N_f = 13 W = 24%	41.6 Ma N_f = 11 W = 17%	57.5 Ma N_f = 21 W = 27%	88.0 Ma N_f = 3 W = 15%	170 Ma N_f = 2 W = 23%	χ^2 = 0.062 ν = 135 ΣN_f = 50.0	P(F,1,ν) = 0% F = 159

Notes: This example was calculated using a distribution of 50 FT zircon grain ages from the Mount Tom sample (ZD6) shown in table 1. Peaks were fit using the Gaussian routine with a separate best-fit relative standard deviation found for each peak. Note that the results shown here are slightly different from those reported for this sample in Brandon and Vance (this issue) because of a change in the number of observations used to represent the composite probability density function.

* Age is the best-fit peak age, N_f = estimated number of grains in the peak, W = estimated relative standard deviation of the peak, and n.d. = peak not detected.

† Symbols for fit statistics are: χ^2 = goodness-of-fit parameter, ν = degrees of freedom, ΣN_f = sum of the estimated number of grains in all fitted peaks, F = F-ratio which provides a measure of the successive improvement in χ^2 with increasing number of best-fit peaks, and P(F, 1, ν) = probability that the F-ratio value is due to chance alone.

My experience indicates that these criteria usually produce a clear-cut solution. Nonetheless, the operator remains the final judge of which solution is both terse and sufficient.

Estimated standard errors for the best-fit parameters.—The Levenberg-Marquardt method produces an estimated covariance matrix of the standard errors for the fitted parameters (Press and others, 1986, p. 525). These standard error estimates should be viewed as rough approximations because of the nonlinear nature of the model and also because the measurement errors may not be normally distributed (see Press and others, 1986, p. 534). Nonetheless, they provide a useful indication of the precision of the best-fit parameters.

Special attention must be given to the standard error of the peak age, because only the variability due to grain error is included in the estimated standard errors from the covariance matrix. Thus, we must add the group error to determine s_{f_j}, the total standard error for the best-fit peak age,

$$s_{f_j} = \sqrt{SE_{fit}^2 + \mu_{f_j}^2 RE_{group}^2} \tag{15}$$

For this equation, SE_{fit} is the estimated standard error from the covariance matrix, and RE_{group} is the "group-only" relative standard error.

CONCEPTUAL BASIS FOR GAUSSIAN PEAK-FITTING METHOD

With respect to the peak-fitting method, there are three questions that need to be addressed: (1) Does the composite probability density plot provide a useful representation of a grain-age distribution? (2) What is the relationship between the parent grain-age distribution and a sample distribution drawn from that parent distribution? and (3) How should the best-fit Gaussian parameters be interpreted?

Justification of composite probability density plot.—The Gaussian peak-fitting method is built around the composite probability density plot. Therefore, it is important to scrutinize more closely the basis of this plot, especially in light of critical comments by Galbraith (1988, p. 277). His discussion implies that the plot is merely an empirical device with no basis in statistical theory.[2] Galbraith (1988) refers to the plot as a "weighted" histogram, but this description is inaccurate, because each grain age is weighted equally. Nor is it a smoothed histogram because smoothing implies a loss of information. The composite probability density plot clearly carries more information than a grain-age histogram.

In my opinion, the composite probability density plot is probably best described as a continuous histogram. For a conventional histogram, the probability mass for each grain age is represented by an incremental increase in the height of the histogram bar that overlies the measured age. In other words, the probability mass is spread as a uniform distribu-

[2] Silverman (1986) provides a comprehensive review of various methods of density estimation. The composite probability density plot of Hurford, Fitch, and Clarke (1984) is a modified version of the kernel method described in Silverman (1986, p. 13–19).

tion over the width of the bar. Thus, a conventional histogram provides a step-wise approximation of the true grain-age frequency distribution. In comparison, the continuous histogram should provide a better approximation because both the grain ages and their associated uncertainties are used. The underlying premise is that the expectation of the measured grain age is more accurately represented by a probability density function rather than a histogram bar of arbitrary width.

To illustrate this point, consider the intermediate case where a grain-age histogram is constructed using a narrow interval, let's say 1 my per interval. Given a typical grain-age distribution, the resulting histogram would be almost useless, because the histogram interval is so small and because the expected number of grain ages per interval is limited to integer values. As a result, there would be little variation among the histogram bars, with almost all the bars limited to a height of either 0 or 1. However, if each grain age is represented by its Gaussian probability density function (eqs 1 and 4 with $\Delta x = 1$ my), then its probability mass would be allocated over a range of histogram intervals. In this case, the bar heights would take on fractional values, so the resulting histogram would more closely approximate the true grain-age frequency distribution. Note that in the limit as Δx approaches zero, this histogram becomes a continuous histogram.

In summary, the continuous histogram provides a useful method for portraying grain-age distributions because: (1) it incorporates more information than a conventional histogram and therefore provides a more complete picture of the probability distribution of a population of grain ages, and (2) it circumvents the problems associated with selecting the size and position of histogram intervals which tend to produce spurious results in conventional histograms.

Galbraith (1988, 1990) has argued that his radial plot provides a better graphical representation of mixed grain-age distributions. Its chief advantage is that it can display both the age and uncertainty of each determination in a group of isotopic ages. There are certainly applications where this feature is important, but in my opinion, the radial plot is not well suited for interpreting mixed grain-age distributions. One problem is that it tends to emphasize individual grain ages, especially the most precise of these ages which plot the farthest from the origin of the plot. As a result, there is very little visual indication of the density of grain ages with changing age, making it difficult to determine the position, separation, and size of potential component grain-age populations. The composite probability density plot is better suited for that purpose.

Sample distribution versus parent distribution.—A *sample* distribution of grain ages provides only an approximate representation of the *parent* distribution from which it was drawn. In this context, it is best to consider peak fitting as a method for decomposing a sample grain-age distribution into a simpler form. The method does provide information about the statistical properties of the sample distribution and its component popu-

lations. However, the limited number of grain ages in a given sample distribution and the even smaller number of grain ages in each component population reduce the precision of these estimates and can also introduce biases.

Of all of the best-fit Gaussian parameters, μ_{fj} is probably least affected by the sampling problem. It should provide a reasonable estimate of the mean age of the jth *parent* population, which comprises all the zircon grain ages in the "jth source terrain." The estimate of the relative standard deviation W_{fj} will be downward biased for component populations with small sample sizes, but this effect is limited because W_{fj} can be no smaller than the quadratic mean of the relative standard errors of the grain ages in the jth component population, as shown in the next section. This lower limit is about 13 percent, given typical "grain-only" relative standard errors.

The sampling problem is mainly manifested in the estimates of N_{fj} and in the detection of component populations that make up a small proportion of the parent distribution. The problem here is that the variable N_{fj} is subject to considerable variability due to geologic factors as well as Poisson-like variations during sampling. To examine this problem, let's restrict ourselves to the second source of variation. The proportion of grain ages in the jth peak of the parent distribution is designated as π_j. In a sample distribution containing N_t grain ages, the expected number of grains for the jth peak should be $N_j = \pi_j N_t$. If $N_t \gg N_j$, then N_{fj}, which is a measured estimate of N_j, should follow a Poisson distribution (note that geologic variability, if present, would add to this Poisson variability). The relative standard error (RE) for a Poisson-distributed variable can be estimated as

$$RE(N_{f_j}) = (N_t \pi_j)^{-1/2} \approx N_{f_j}^{-1/2} \tag{16}$$

This relationship shows that when the expected value for N_{fj} is small, the measured values for N_{fj} will vary widely due to this Poisson sampling effect. For example, $N_{fj} = 10$ has $RE(N_{fj}) = \pm 32$ percent, and $N_{fj} = 5$ has $RE(N_{fj}) = \pm 45$ percent.

This situation brings up a related problem, that is the detection of small peaks. The concept of a detection limit is commonly used in the analysis of trace concentrations of a chemical element (Potts, 1987). In this context, the detection limit is defined as the concentration at which the estimated relative standard error of a measurement method exceeds 33 percent. The rationale is that at the detection limit, the measured concentration is indistinguishable from zero because the 99 percent confidence limit, which is delimited by $\pm 3RE = \pm 100$ percent around the measured concentration, overlaps with zero.

Using eq (16), this concept can be adopted to define π_{DL}, which marks the lower limit of confident detection of a component population that makes up a specified proportion, π_j, of the parent distribution. This proportion is said to be below the detection limit when $\pi_j < \pi_{DL}$.

$$\pi_{DL} = \frac{9}{N_t} \tag{17}$$

Using typical values for N_t of 25, 50, and 75 grain ages, eq (17) predicts π_{DL} is 36, 18, and 12 percent, respectively. Experience indicates that FT grain-age distributions commonly contain 3 to 5 component populations (table 2), with each of these populations making up, on the average, about 33 to 20 percent of the total distribution. The detection limit concept indicates that the smaller peaks, which will have proportions less than average, might be easily missed in a typical sample where $N_t \approx 50$.

Interpretation of best-fit Gaussian parameters.—At this point, it is important to examine more closely the relationship between the best-fit Gaussian parameters μ_{fj} and W_{fj} and the statistical properties of the jth component population in a sample grain-age distribution. To illustrate, consider a simple example provided by zircon grain ages from the Fish Canyon tuff (fig. 4). The single peak in the composite probability density plot is expected since these zircons were derived from a single-age volcanic source. For this example, g = 1 so that the following relationships hold: j = 1, $N_{fj} = N_t$, $p_{fj}(x) = p_t(x)$, and $\mu_{fj} = \mu_t$. The mean of this distribution can be calculated directly from the composite probability

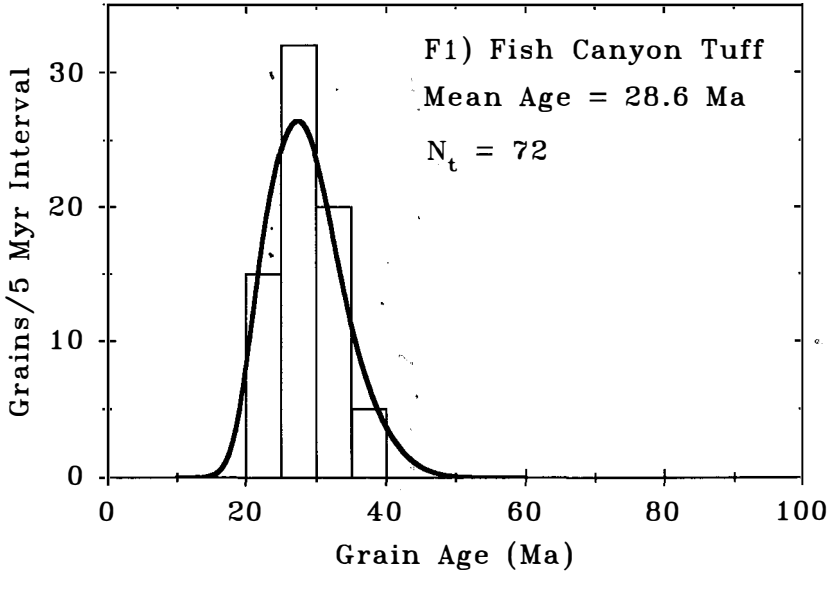

Fig. 4. Composite probability density plot and grain-age histogram for zircons from the Fish Canyon tuff (sample F1 in Brandon and Vance, this issue), which serves as a good example of a grain-age distribution derived from a single-age source.

density function (1,2). The formula for the mean of a normal distribution gives the following result:

$$\mu_{f_j} = \int_{-\infty}^{+\infty} x\, p_{f_j}(x)dx = \frac{1}{N_{f_j}} \int_{-\infty}^{+\infty} x \sum_{i=1}^{N_{f_j}} p_i(x)dx = \frac{1}{N_{f_j}} \sum_{i=1}^{N_{f_j}} \int_{-\infty}^{+\infty} x\, p_i(x)dx$$

$$\mu_{f_j} = \frac{1}{N_{f_j}} \sum_{i=1}^{N_{f_j}} \mu_i \tag{18}$$

As might be expected, the best-fit Gaussian mean μ_{fj} is equivalent to the mean of the grain ages μ_i.

But what about W_{fj}? For this case, let's examine the related parameter σ_{fj}, the absolute standard deviation of the best-fit Gaussian, where $\sigma_{fj} = \mu_{fj} W_{fj}$. Once again using the single peak example (g = 1), the formula for the absolute standard deviation of a normal distribution gives the following result:

$$\sigma_{f_j}^2 = \int_{-\infty}^{+\infty} (x - \mu_{f_j})^2 p_{f_j}(x)dx = \frac{1}{N_{f_j}} \int_{-\infty}^{+\infty} (x - \mu_{f_j})^2 \sum_{i=1}^{N_{f_j}} p_i(x)dx$$

$$= \frac{1}{N_{f_j}} \sum_{i=1}^{N_{f_j}} \int_{-\infty}^{+\infty} (x - \mu_{f_j})^2 p_i(x)dx$$

This equation is modified by substituting $(x - \mu_i) + (\mu_i - \mu_{fj})$ for $(x - \mu_{fj})$ and multiplying out the squared term, which gives

$$\sigma_{f_j}^2 = \frac{1}{N_{f_j}} \sum_{i=1}^{N_{f_j}} \int_{-\infty}^{+\infty} [(\mu_i - \mu_{f_j})^2 + (x - \mu_i)^2 + 2(\mu_i - \mu_{f_j})(x - \mu_i)] p_i(x)dx$$

The third term in the square brackets drops out because the integral of that term is zero. This leaves only the first two terms which can be reduced to the following form

$$\sigma_{f_j}^2 = SD_{f_j}^2 + \frac{1}{N_{f_j}} \sum_{i=1}^{N_{f_j}} s_i^2 \tag{19}$$

Eq (19) shows that σ_{fj}^2 is the sum of two variance terms. The first is SD_{fj}^2, the variance of the grain ages relative to the estimated mean age of the peak. The second is the average variance due to the standard errors of the grain ages (s_i). This relationship illustrates the distinction between the best-fit Gaussian parameter W_{fj}, which measures the relative standard deviation of a peak in the composite probability density plot, and SD_{fj}/μ_{fj}, which represents the relative standard deviation of the grain ages in the jth component population. Even if W_{fj} is not a direct measure of the grain-age relative standard deviation, it does provide a useful index of the variability among grain ages within a component population.

W_{fj} measures the total random variability of the grain ages in the jth sample population. One important source of variability is the Poisson process of radioactive decay. There are other sources, however, such as geologic factors and random laboratory errors that can contribute to W_{fj} (see Green, 1981, p. 80–81, for a discussion of non-Poisson errors in FT dating). Possible geologic factors are differential erosion, prolonged igneous activity, slow uplift, or compositional variations between zircon grains which might affect annealing or etching of the zircons.

The decomposition method of Galbraith and Green (1990) considers the variation among the grain ages in a component population to be solely a function of the Poisson process of radioactive decay. In contrast, W_{fj} in the Gaussian peak-fitting method is retained as a set of free parameters so that the contribution of non-Poisson errors can be directly assessed. A population of zircons derived from a single-age FT source is estimated to have $W \approx 16.3$ percent as indicated by FT age determinations for unreset volcanic zircons from two tuff samples reported in Brandon and Vance (this issue, samples C5 and F1). This reference value can be compared to estimated values of W_{fj}, as determined by the Gaussian peak-fitting method, to assess whether or not an identified FT grain-age peak is compatible with a single-age FT source. In the study of Brandon and Vance (this issue), which examines FT grain-age distributions from 15 unreset sandstone samples, many of the identified peaks had W_{fj} values of about 17 to 18 percent, which suggests they were derived from single-age FT source terrains. However, a significant proportion of the peaks, especially the older peaks, had W_{fj} values of 20 to 33 percent, suggesting they were derived from more internally complex FT source terrains.

The analysis above only considers the case of a single peak. For the more complex case of a finite-mixture distribution, it is inferred that μ_{fj} and W_{fj} are reasonable estimates of the Gaussian parameters for the jth sample population, even though we cannot specifically identify all the grains in that component population. Clearly, there are cases where the best-fit Gaussian parameters fail to represent the underlying grain-age populations. Probably the most common situation is when two Gaussians are so closely spaced that they cannot be resolved by the fitting algorithm. The limiting factor in this case is usually the size of the grain age uncertainties s_i, not the intrinsic capabilities of the fitting routine. The next section more closely examines this problem of peak resolution.

TESTING THE DECOMPOSITION METHODS

Five artificial binary mixtures were constructed using FT grain-age distributions for unreset volcanic zircons from two tuff samples reported in Brandon and Vance (this issue, samples C5 and F1). Each of these grain-age distributions is considered to represent a typical population of FT grain-ages derived from a single-age source. Table 3 summarizes the mixture proportions and expected ages for each of the component

TABLE 3

Test of χ^2 age and Gaussian peak-fitting methods using artificial binary mixtures of FT zircon grain ages

Proportions, Separation (Δ)	Age Range, Number of Grains	χ^2 age ± 2s	Young Population			Old Population		
			Expected*	Fit #1*	Fit #2*	Expected*	Fit #1*	Fit #2*
A) 50% 30 Ma 50% 51 Ma Δ = 3.7	22 - 66 Ma N_t = 40	29.7 ± 2.2 Ma N_f = 20	30.0 Ma N_f = 20 W = 16%	29.4 Ma N_f = 18 W = 16%	29.7 Ma N_f = 19 W = 16%	51.1 Ma N_f = 20 W = 15%	49.7 Ma N_f = 22 W = 18%	50.4 Ma N_f = 21 W = 16%
B) 50% 30 Ma 50% 44 Ma Δ = 2.6	22 - 57 Ma N_t = 40	30.2 ± 2.2 Ma N_f = 22	30.0 Ma N_f = 20 W = 16%	29.3 Ma N_f = 15 W = 15%	30.1 Ma N_f = 19 W = 17%	44.0 Ma N_f = 20 W = 15%	41.7 Ma N_f = 25 W = 20%	43.5 Ma N_f = 21 W = 17%
C) 26% 29 Ma 74% 44 Ma Δ = 2.8	22 - 57 Ma N_t = 27	29.3 ± 2.7 Ma N_f = 8	29.0 Ma N_f = 7 W = 16%	26.6 Ma N_f = 3 W = 14%	28.9 Ma N_f = 6 W = 17%	44.0 Ma N_f = 20 W = 15%	42.0 Ma N_f = 24 W = 19%	43.3 Ma N_f = 21 W = 17%
D) 50% 30 Ma 50% 39 Ma Δ = 1.8	22 - 51 Ma N_t = 40	31.1 ± 2.2 Ma N_f = 29	30.0 Ma N_f = 20 W = 16%	30.5 Ma N_f = 19 W = 16%	30.9 Ma N_f = 23 W = 17%	39.3 Ma N_f = 20 W = 15%	38.5 Ma N_f = 21 W = 18%	39.8 Ma N_f = 17 W = 17%
E) 33% 30 Ma 66% 39 Ma Δ = 1.7	25 - 51 Ma N_t = 30	34.3 ± 2.5 Ma N_f = 26	30.2 Ma N_f = 10 W = 16%	29.6 Ma N_f = 6 W = 14%	31.3 Ma N_f = 13 W = 16%	39.3 Ma N_f = 20 W = 15%	38.2 Ma N_f = 23 W = 18%	40.0 Ma N_f = 17 W = 16%

Notes: N_t = total number of grains analyzed. N_f = estimated number of grains in a specific peak or fraction. All uncertainties are cited at ± 2 standard error. W = estimated relative standard deviation of a peak.

* Expected values for each peak were determined by finding best-fit Gaussian parameters for each component population. The binary mixtures were decomposed using two different models. For fit #1, a separate best-fit relative standard deviation was found for each peak. For fit #2, a common best-fit relative standard deviation was calculated for both peaks in the distribution.

populations, and the decomposition results as determined by the Gaussian peak-fitting method.

Each of the component populations was generated by random selection from one of the "parent populations," until the required number of grain ages was obtained. The younger component population has an expected peak age of about 30 Ma. The older component population has three different expected peak ages: 39.3, 44.0, and 51.1 Ma. This variation was introduced to examine the effect of peak separation on the performance of the decomposition methods. The different peak ages were produced by multiplying the spontaneous track count for each grain by a constant. The final step was to determine the χ^2 age and best-fit peak ages for each of the binary mixtures. Composite probability density plots are shown in figure 5.

As shown in table 3, the χ^2 method and the Gaussian peak-fitting method both produce good estimates of the expected Gaussian parameters of the component populations. Experiment A is the easiest to analyze; the composite probability density plot (fig. 5A) shows an obvious bimodal pattern. The peak ages for experiment B are more closely spaced; the resulting composite probability density plot (fig. 5B) is unimodal. The small shoulder to the right of the mode is the only visible evidence of a second grain-age population. For experiment C, the composite probability density plot (fig. 5C) is also unimodal. The only visible evidence of a second population is the long tail to the left of the mode. For experiment D, the composite probability density plot (fig. 5D) shows no obvious evidence of two populations other than a slight negative kurtosis. However, it is this departure from an ideal Gaussian shape that allows the peak-fitting routine to identify correctly the component populations. Experiment E is the hardest to decompose. The composite probability density plot (fig. 5E) looks like a nearly perfect Gaussian. Even so, the peak-fitting routine gives a reasonable result, but the χ^2 age seriously overestimates the age of the young peak.

Two different results, labeled fit #1 and fit #2, are given for the Gaussian peak-fitting method. For fit #1, a separate best-fit relative standard deviation (W_{fj}) was found for each peak. For fit #2, a common best-fit relative standard deviation was calculated for both peaks of the distribution. The rationale behind this second approach is that if all peaks are derived from single-age sources, then W_{fj} should be approximately the same for all peaks. Because this approach uses a smaller set of more highly constrained parameters, it should result in a more stable and better resolved solution, so long as the assumption of single-age peaks remains a reasonable one, as it is for the test cases in table 3. Note that fit #2 generally does give a solution closer to the expected values. There may be real situations where one might be able to justify the use of this type of model structure.

Generalizations.—The above experiments illustrate some general features of the χ^2 and Gaussian peak-fitting methods. First, it is clear that the two methods can produce quite different results, as shown in experiment E.

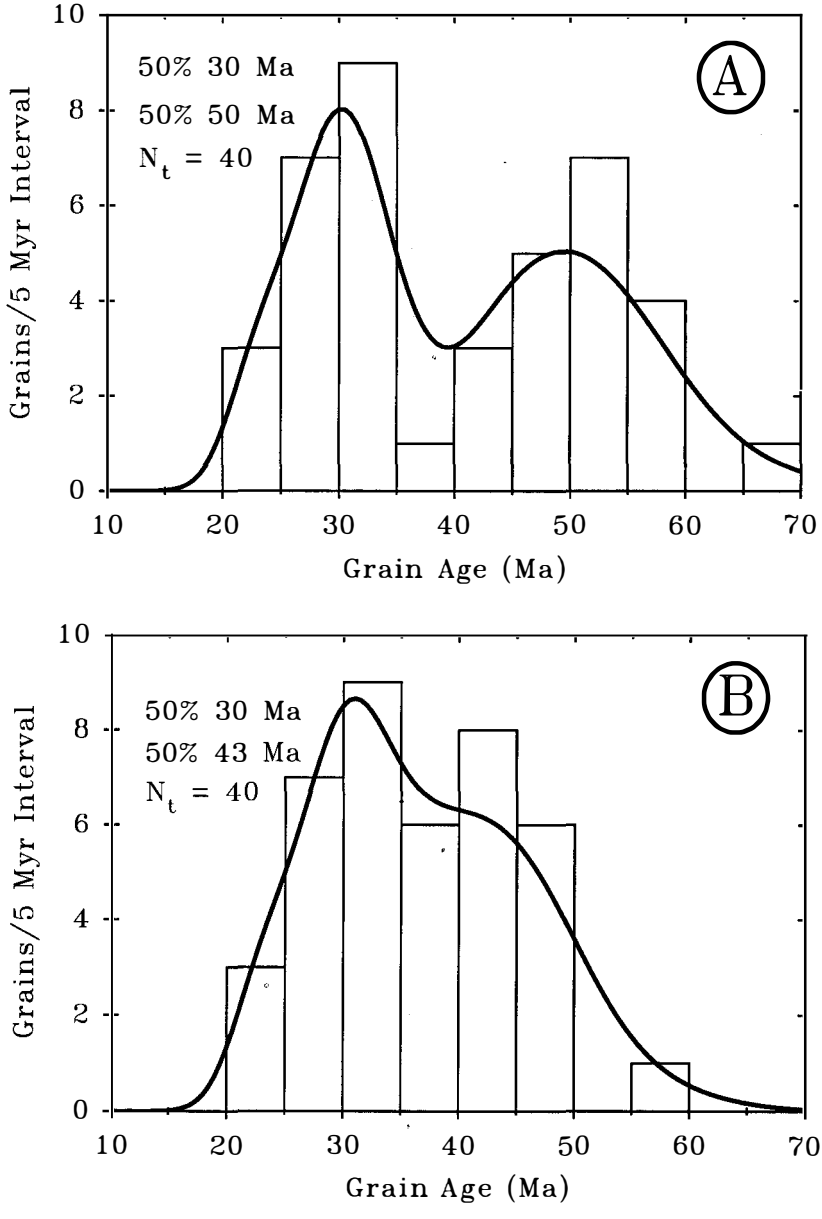

Fig. 5(A–E). Composite probability density plots and grain-age histograms for simulated binary mixtures of FT zircon grain ages. These grain-age distributions are used to test the decomposition methods as summarized in tables 3 and 4. See text for further details.

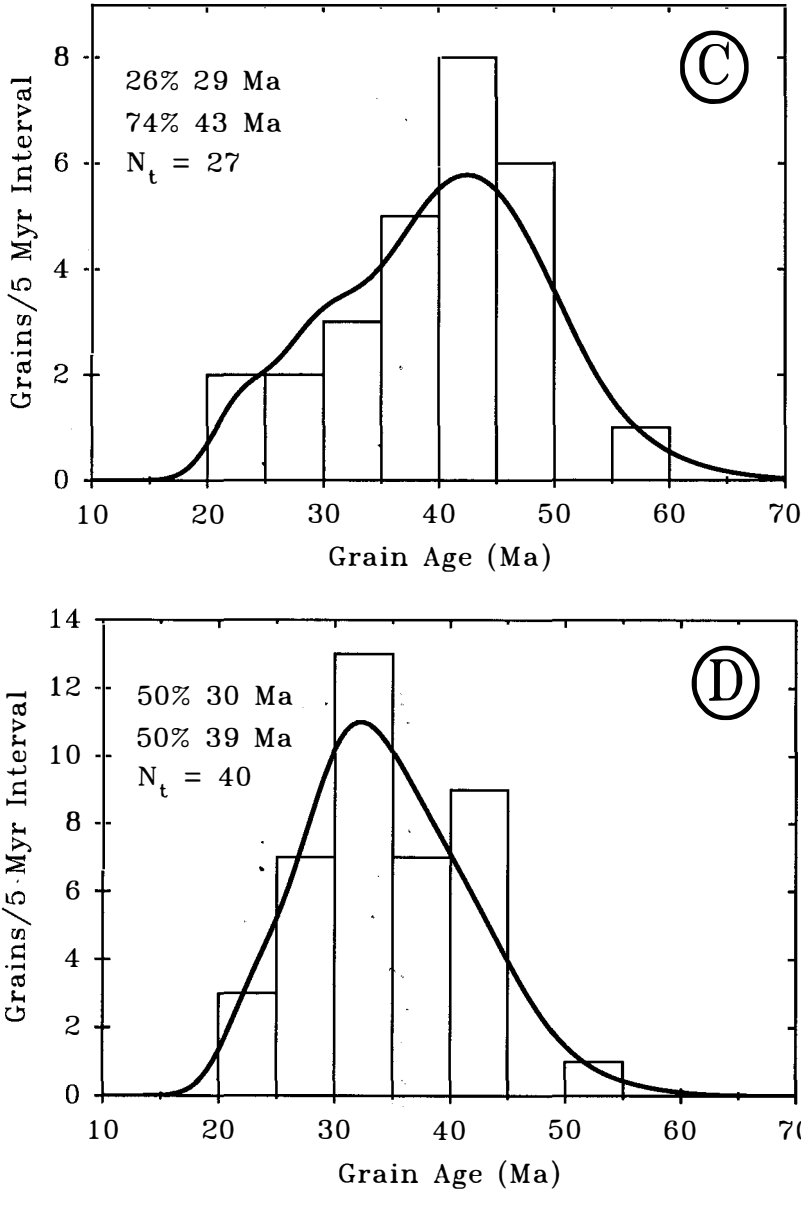

Fig. 5(continued)

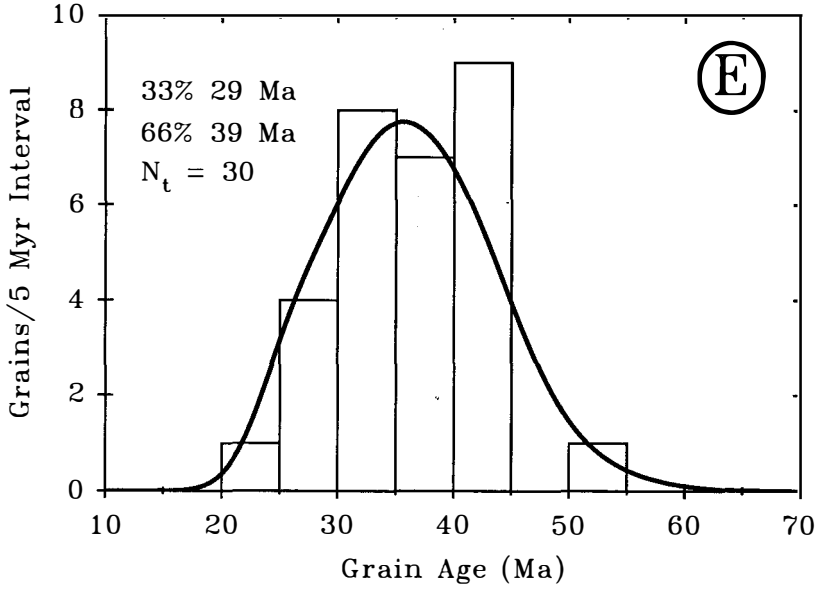

Fig. 5(continued)

But this reflects the different approach used by them. The peak-fitting method fits component Gaussians to the main bumps and shoulders in the composite probability density plot. It provides little information about the statistical significance of these features. In contrast, the χ^2 age method isolates the largest group of young grain ages for which there still exists a reasonable probability of association. This method is entirely blind to the bumps and shoulders in the composite probability density plot, but it does contain a measure of statistical significance.

In general, these two methods are usually complementary, especially if the main task is to establish a maximum limit for the depositional age of an unreset sandstone sample. In this case, the χ^2 age method is preferred because any biases in the age estimate will tend to be conservative, that is, the age assigned to the youngest fraction of grain ages may be slightly too old. Even so, the peak-fitting method has its place in that it provides additional information that can be used to judge the quality of this estimated age. Consider the three possible results: (1) The most optimal result is when there is close agreement between the χ^2 age and the youngest best-fit peak age. In this case, the youngest fraction of grains is distinguished both as a unique Gaussian peak and as a statistically distinct component of the overall distribution. (2) A less optimal result is when the χ^2 age is significantly younger than the youngest best-fit peak. Experiment D might have produced this result if the two component populations were more closely spaced. In this case, the χ^2 age method

will, once again, isolate a statistically distinct fraction of the overall distribution; however, the possibility remains that this fraction represents only part of the youngest component population. In other words, the source of the youngest zircon population may not be restricted to a "single age". (3) The least optimal result is when the χ^2 age is significantly older than the youngest best-fit peak. A good example is experiment E. A best-fit youngest peak has been identified, but there is no assurance that the peak includes a statistically distinct group of grain ages. The χ^2 age still provides a maximum age for the youngest fraction of grain ages in the sample; however, the possibility remains that the parent distribution contains a distinct component population that is younger than the χ^2 age. More grain ages would be needed to resolve this possibility.

Peak separation.—The experiments in table 3 show the important effect that peak separation has on the performance of the decomposition methods. General aspects of this problem are discussed in Everitt and Hand (1981, p. 27–30) and Titterington, Smith, and Makov (1985, p. 48). A useful measure of the degree of separation between two peaks is provided by

$$\Delta = \left[\frac{(\mu_{f_1} - \mu_{f_2})^2 (\sigma_{f_1}^2 + \sigma_{f_2}^2)}{2\sigma_{f_1}^2 \sigma_{f_2}^2} \right]^{1/2} \tag{20}$$

which is modified from Everitt and Hand (1981, eq 2.4). Note that absolute standard deviations are used here instead of relative standard deviations ($\sigma_{f1} = \mu_{f1} W_{f1}$). When $\sigma_{f1} = \sigma_{f2}$, eq (20) takes on a simpler form,

$$\Delta = \frac{|\mu_{f_1} - \mu_{f_2}|}{\sigma_f} \tag{21}$$

Thus, Δ represents the distance between the peaks as normalized to some measure of peak width. Table 3 shows the Δ values calculated for the five artificial mixtures using eq (20). These examples indicate some guidelines useful in similar cases where two component populations make up a significant proportion of the total distribution ($\pi_j > \sim 20$ percent) and are well separated from other component populations. (1) When $\Delta > \sim 3.5$, the two components will appear as two distinct modes in the composite probability density plot (fig. 5A). (2) When $\sim 3.5 > \Delta > \sim 2.0$, the composite probability density plot will be bitangential (fig. 5B). (3) When $\Delta = \sim 1.7$, the composite probability density plot approaches the form of a single ideal Gaussian (fig. 5E). In this case, decomposition of the two component populations might be difficult. For component populations derived from single-age sources ($W \approx 16$ percent), we can use eq (20) to estimate the lower resolution limit for separating two significant peaks, defined as when $\Delta < 1.7$. In this case, the peak ages must be separate by more than about 7 my for peaks centered around 25 Ma, 13 my for 50 Ma, and 25 my for 100 Ma.

Binomial peak-fitting method.—The binomial peak-fitting method of Galbraith and Green (1990) was also tested using the artificial binary mixtures described above. Results are show in table 4. This method works very well and in fact may be the preferred peak-fitting method when there are only 2 or 3 peaks in the distribution. Galbraith and Green's (1990) method is based on the maximum likelihood approach, which is a statistically elegant approach especially as applied to the problem of mixed FT grain-age distributions. One potential disadvantage of their method is that it assumes that variations among grain ages in a component population are due solely to the Poisson process of radioactive decay. This assumption may be too restrictive given that geologic factors might introduce an additional source of variability. Another disadvantage is that there is little known about the performance of their binomial peak-fitting method with grain-age distributions containing more than 2 peaks. My limited experience suggests that the method has difficulties when it is used to fit 4 or more peaks. Finding a best-fit solution for their binomial peak-fitting method involves a non-linear search in parameter space for the global maximum of the likelihood surface. Galbraith and Green (1990) use the EM algorithm (Titterington, Smith, and Makov, 1985, p. 84) to find the best-fit solution, which is formally called the maximum likelihood solution. The Gaussian peak-fit routine, introduced in this paper, also conducts a non-linear search, although this search attempts to find a global minima in the χ^2 surface. Titterington, Smith, and Makov (1985, p. 82–105) discuss the relative merits and problems of these two optimization methods when applied to the problem of decomposition of finite-mixture distributions. Their review suggests that the maximum likelihood method using the EM search algorithm might be more prone, under certain circumstances, to halt at local maxima on the likelihood surface, thus failing to converge to the maximum likelihood solution. One gets the impression, however, that a detailed comparison between these two methods has not yet been done. Further investigation is needed before it can be decided which peak-fitting method is best for the decomposition of FT grain-age distributions.

CONCLUSIONS

FT dating of detrital zircons from unreset sandstones has much potential for resolving important problems concerning the age and provenance of epiclastic rocks. The methods introduced here provide a means for decomposing a large distribution of grain ages into a simpler and more meaningful form. The χ^2 age method provides a straightforward method for estimating the pooled age of the youngest fraction of FT grain ages in a sample grain-age distribution. It is particularly useful for providing a maximum limit for the depositional age of an unreset sample. The Gaussian peak-fitting method can be used to estimate the age, relative standard deviation, and size of various component grain-age populations in a sample distribution. This information can be used to

TABLE 4

Test of the binomial peak-fitting method of Galbraith and Green (1990) using artificial binary mixtures of FT zircon grain ages

	Proportions in mixture	Age Range, Number of Grains	Young Population		Old Population	
			Expected*	Fit #3*	Expected*	Fit #3*
A)	50% 30 Ma 50% 51 Ma	22 - 66 Ma $N_t = 40$	29.7 ± 2.2 Ma $N_f = 20$	29.6 ± 2.2 Ma $N_f = 20$	50.9 ± 3.4 Ma $N_f = 20$	50.8 ± 3.9 Ma $N_f = 20$
B)	50% 30 Ma 50% 44 Ma	22 - 57 Ma $N_t = 40$	29.7 ± 2.2 Ma $N_f = 20$	29.6 ± 2.4 Ma $N_f = 19$	43.9 ± 3.0 Ma $N_f = 20$	43.5 ± 3.7 Ma $N_f = 21$
C)	26% 29 Ma 74% 44 Ma	22 - 57 Ma $N_t = 27$	28.4 ± 2.8 Ma $N_f = 7$	28.2 ± 3.1 Ma $N_f = 7$	43.9 ± 3.0 Ma $N_f = 20$	43.6 ± 3.5 Ma $N_f = 20$
D)	50% 30 Ma 50% 39 Ma	22 - 51 Ma $N_t = 40$	29.7 ± 2.2 Ma $N_f = 20$	29.0 ± 3.7 Ma $N_f = 16$	39.2 ± 2.7 Ma $N_f = 20$	38.0 ± 4.4 Ma $N_f = 24$
E)	33% 30 Ma 66% 39 Ma	25 - 51 Ma $N_t = 30$	30.0 ± 2.7 Ma $N_f = 10$	29.4 ± 4.5 Ma $N_f = 9$	39.2 ± 2.7 Ma $N_f = 20$	38.8 ± 3.8 Ma $N_f = 21$

Notes: N_t = total number of grains analyzed. N_f = estimated number of grains in a specific peak. Ages correspond to peak ages with uncertainties of ± 2 standard error.

* Expected values for each component population were determined by fitting a single peak to the component population using the binomial peak-fitting method. Fit #3 marks the best-fit peaks for the binary mixtures as determined using the binomial peak-fitting method.

identify specific FT source terrains in the source region of a sedimentary basin and to track the path of sediment transport from source to basin.

Finally, it is important to note that estimation and inference testing for finite-mixture distributions remain a challenging and incompletely resolved problem in statistical research. As a result, there is much to be learned from practical application of decomposition methods using real FT grain-age distributions. Any user of the methods advocated here and those advocated by Galbraith and Green (1990) should remain watchful for potential problems, especially if the number of samples being analyzed is small. Perhaps the most useful test of these methods will be their ability to resolve credible information from large groups of samples.

ACKNOWLEDGMENTS

I am greatly thankful to Joseph Vance for initiating my interest in using the FT method for dating sedimentary units, for his enthusiastic collaboration on our FT dating project in the Olympic subduction complex, and for his patient advice about the practical side of FT dating. Discussions with Sean Willett and Jeff Park resulted in some important improvements to the peak-fitting method and helped sharpen my understanding of the composite probability density plot. The paper was much improved by critical reviews by Mark Cloos, John Garver, Rowland Tabor, Joseph Vance, and Sean Willett. Note, however, that I accept full responsibility for the interpretations and arguments advanced herein. Research funding from the National Science Foundation (EAR-8707442, EAR-9005777) and Yale University is gratefully acknowledged.

REFERENCES

Baldwin, S. L., Harrison, T. M., and Burke, K., 1986, Fission-track evidence for the source of accreted sandstones, Barbados: Tectonics, v. 5, p. 457–468.

Bevington, P. R., 1969, Data reduction and error analysis for the Physical Sciences: New York, McGraw-Hill Book Company, 336 p.

Cerveny, P. F., Naeser, N. D., Zeitler, P. K., Naeser, C. W. and Johnson, N. M., 1988, History of uplift and relief of the Himalaya during the past 18 million years: Evidence from fission-track ages of detrital zircons from sandstones of the Siwalik group, *in* Kleinspehn, K. L., and Paola, C., editors, New Perspectives in Basin Analysis: New York, Springer-Verlag, p. 43–61.

Everitt, B. S., and Hand, D. J., 1981, Finite mixture distributions: London, England, Chapman and Hall, 143 p.

Galbraith, R. F., 1981, On statistical models for fission-track counts: Journal of Mathematical Geology, v. 13, p. 471–478.

———— 1988, Graphical display of estimates having differing standard errors: Technometrics, v. 30, p. 271–281.

———— 1990, The radial plot: graphical assessment of spread in ages: Nuclear Tracks and Radiation Measurements, v. 17, p. 207–214.

Galbraith, R. F., and Green, P. F., 1990, Estimating the component ages in a finite mixture: Nuclear Tracks and Radiation Measurements, v. 17, p. 197–206.

Gleadow, A. J. W., 1981, Fission-track dating methods: what are the real alternatives?: Nuclear Tracks, v. 5, p. 3–14.

Green, P. F., 1981, A new look at statistics in fission-track dating: Nuclear Tracks, v. 5, p. 77–86.

Hurford, A. J., 1990, Standardization of fission track dating calibration: Recommendation by the Fission Track Working Group of the I.U.G.S. Subcommission on Geochronology: Chemical Geology (Isotope Geoscience Section) v. 80, p. 171–178.

Hurford, A. J., Fitch, F. J., and Clarke, A., 1984, Resolution of the age structure of the detrital zircon populations of two Lower Cretaceous sandstones from the Weald of England by fission track dating: Geological Magazine, v. 121, p. 269–277.

Kowallis, B. J., Heaton, J. S., and Bringhurst, K., 1986, Fission-track dating of volcanically derived sedimentary rocks: Geology, v. 14, p. 19–22.
McLachlan, G. J., and Basford, K. E., 1988, Mixture models: New York, Marcel Dekker, Incorporated, 253 p.
Menke, W., 1989, Geophysical data analysis: discrete inverse theory (revised edition): San Diego, California, Academic Press, Incorporated, 289 p.
Naeser, N.D., Naeser, C. W., and McCulloh, T. H., 1989, The application of fission-track dating to the depositional and thermal history of rocks in sedimentary basins, *in* Naeser, N. D., and McCulloh, T. H., editors, Thermal history of sedimentary basins: New York, Springer-Verlag, p. 157–180.
Potts, P. J., 1987, A handbook of silicate rock analysis: New York, Chapman and Hall, 622 p.
Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., 1986, Numerical recipes: Cambridge, England, Cambridge University Press, 818 p.
Seward, D., and Rhoades, D. A., 1986, A clustering technique for fission-track dating of fully to partially annealed minerals and other non-unique populations: Nuclear Tracks and Radiation Measurements, v. 11, p. 259–268.
Silverman, B. W., 1986, Density estimation for statistics and data analysis: New York, Chapman and Hall, 175 p.
Taylor, J. R., 1982, An introduction to error analysis: Mill Valley, California, University Science Books, Oxford University Press, 270 p.
Titterington, D. M., Smith, A. F. M., and Makov, U. E., 1985, Statistical analysis of finite mixture distributions: New York, John Wiley and Sons, 243 p.