

# Detrital zircons as tracers of sedimentary provenance: limiting conditions from statistics and numerical simulation

Tom Andersen

*Department of Geosciences, University of Oslo, P.O. Box 1047, Blindern, Oslo N-0654, Norway*

Received 8 March 2004; accepted 17 November 2004

---

## Abstract

U–Pb ages of detrital zircons in clastic sediments are potential indicators of sedimentary provenance and crustal evolution. To extract geologically significant information from such data, it is necessary that the zircons analysed in the laboratory reflect the sediment from which they were separated. Among other factors, this depends on the number of grains analysed. The probability that a detrital zircon in a sediment belongs to one of several populations (defined by their age or another well-defined parameter) is described by the binomial probability formula, from which detection limits and expected values for the accuracy and precision of population sizes can be derived. Monte Carlo simulation of sets of data with up to 120 individual analyses drawn from pools with known age probability density distributions shows that quantitative representation of the age distribution of the zircons in the host sediment is unlikely to be achieved in natural cases. Methods that assume such representation (e.g., age spectrum deconvolution, multivariate statistical analysis using probability density scores, and mass balance modelling) are therefore likely to give results of questionable significance. The best way to extract the potential information contained in detrital zircons may be a combination of random and nonrandom selection of grains for analysis; yielding two complementary sets of age data from each sample, the random fraction should comprise 35–70 grains or more, depending on the complexity of the age distribution. The use of trace element and Lu–Hf isotope data from the zircons dated by U–Pb will give additional, valuable information, which may facilitate interpretation of detrital zircon ages. The use of detrital zircon ages to limit the age of deposition of the host sediment is fraught with both geological and statistical sources of error, and is therefore discouraged.

© 2004 Elsevier B.V. All rights reserved.

*Keywords:* Clastic sediments; Provenance determination; U–Pb; Detritus; Zircon; Statistics

---

## 1. Introduction

U–Pb dating of detrital zircons from clastic (meta) sediments by TIMS-ID, SIMS, and LAM-ICPMS has

become a popular method in sedimentary correlation and provenance studies (Fedó et al., 2003 and references therein). U–Pb ages on single detrital zircon grains are used to identify provenance components in a sedimentary unit (e.g., Haas et al., 1999), to correlate between sequences (e.g., Bingen et al.,

---

*E-mail address:* tom.andersen@geo.uio.no.

2001), to determine a maximum limit for the age of deposition (e.g., Knudsen et al., 1997; Williams, 2001; Bingen et al., 2001), and to study crustal evolution processes on a continentwide scale (e.g., Davis, 2002; Goodge et al., 2002; Barr et al., 2003; van Wyck and Williams, 2002; Griffin et al., 2004).

Most studies published so far have based their conclusions on visual comparison of U–Pb concordia plots or probability density diagrams (Sircombe, 2000) for multigrain sets of single zircons. More quantitative approaches, such as age spectrum deconvolution (Sambridge and Compston, 1994), principal component analysis (Sircombe, 2000), and kernel density estimations (Sircombe and Hazelton, 2004), have also been attempted. However, for any of these methods to give results of geological significance, it is necessary that the age distribution of zircons analysed in the laboratory reflects the age distribution of detrital zircons in the sediment. The geological information that can be extracted from a set of age data on detrital zircons is critically dependent on the number of grains analysed, and on how the selection of grains for analysis is managed. In some applications of detrital zircon data, such as the search for minor exotic components, or the use of detrital zircon ages to limit the age of deposition of a sediment, it is important to detect zircon populations that have low abundances in the sediment. Although the need to analyse statistically significant numbers of zircons has been clearly demonstrated (Dodson et al., 1988; Sircombe, 2000; Vermeesch, 2004), many studies continue to base far-reaching geological conclusions on limited numbers of analyses (e.g., Haas et al., 1999; Dahlgren and Corfu, 2001, Davis, 2002; Barr et al., 2003).

In the present paper, some theoretical, limiting conditions for the use of detrital zircon ages as indicators of sedimentary provenance will be examined from some elementary statistical concepts, and from numerical modeling based on real and synthetic data. The main question to be answered is: When, and to what extent, does a zircon separate reflect the population distribution of its host sediment? Specific problems to be addressed include the significance of statistically defined detection limits (Dodson et al., 1988; Vermeesch, 2004), possible sources of bias in zircon age data, the use of detrital zircons to define limits for the age of deposition of sediments, and the potential of detrital zircon as a mass balance indicator.

To provide examples of the approach, the insights gained will be used to reevaluate some published detrital zircon data from Proterozoic and early Phanerozoic sedimentary rocks from the SW part of the Baltic Shield.

## 2. Statistical relationships and detrital zircon ages

### 2.1. Assumptions and definitions

In the following, lower-case symbols refer to properties of a sample of detrital zircons selected for analysis, whereas upper-case symbols refer to properties of the sediment from which the zircons have been separated. As the total number of zircons analysed in a dating study ( $n$ ) will always be very much smaller than the total number of detrital zircon grains contained in the rock unit sampled ( $N$ ), a sedimentary unit may be seen as an infinite reservoir of detrital zircons. These zircons originally formed by igneous or metamorphic crystallization processes within one or more *protosources* (Pell et al., 1997). It is assumed that each of the  $N$  detrital zircons in a sediment may be unambiguously assigned to one of several *populations*, defined by U–Pb age, Hf isotopic composition, or any other quantitative physical or chemical parameter, with the  $i$ th population comprising  $N_i$  zircons. As most quantitative detrital zircon studies published so far have been concerned with U–Pb ages, only *age populations* will be considered here. An age population simply consists of zircons whose U–Pb ages lie between arbitrarily defined upper and lower age limits. In some studies, *provenance component* is used in this sense (e.g., Sircombe, 2000). A provenance component should ideally consist of material derived from a single protosource or a group of related protosources making up a *provenance terrain*. As each age population may be made up of zircons derived from several unrelated protosources, this usage is unfortunate, and the term *provenance component* will not be used in this study. In the following analysis of zircon populations, it is assumed that the separation of zircons from their host sediment and selection of grains for analysis are strictly random processes (the need for nonrandom sampling, and its consequences, will be discussed separately; see Section 5.4.1).

One of the main problems in detrital zircon geochronology is to determine the value of  $n$  needed to ensure that enough data are collected to detect all populations present. In the present paper, this will be approached through an analysis of the *detection limit* and the *representativity* of a data set. The detection limit is defined as the abundance of the largest population of zircons likely to remain undetected in  $n$  analyses, with no assumptions being made on the number or characteristics of populations actually detected (Fig. 1a). This criterion was introduced by Dodson et al. (1988), and has been used in most subsequent studies (e.g., Fedo et al., 2003).

Vermeesch (2004) proposed an alternative approach to this problem, in which the detection limit is defined as the size of  $m$  equally abundant populations, one of which is likely to escape detection in  $n$  analyses (Fig. 1b). The detection limit defined by this method is larger than by the traditional method (or, using the language of Vermeesch (2004), a larger number of zircons is needed to acquire a desired level of statistical adequacy). However, the usefulness of this estimate is reduced by the need of making assumptions on the nature and distribution of the detected populations. The significance of a zircon population, or “age fraction” in

the terms of Vermeesch (2004), is reduced to that of the bin width in an abundance histogram. The recommended critical number of zircons (117) was based on an assumption of 20 equally abundant (5%) populations. The equal abundance assumption may be useful when deriving a “worst-case estimate,” but it strongly reduces the predictive power of the method. As was also demonstrated by Vermeesch (2004), the detection limit in real cases, where populations are neither equally abundant nor equally spaced, would be somewhere between the equal population limit and the detection limit based on the assumptions of Dodson et al. (1988). It will be shown below that the detection limit is only one of several parameters dependent on  $n$ , which have to be considered when detrital zircon ages are used for geological interpretations; in many cases the detection limit will not be the most critical among these.

## 2.2. Binomial probability

As long as no deliberate or accidental bias is introduced (Sircombe and Stern, 2002), the selection of  $n$  zircons for analysis may be regarded as a series of  $n$  independent, random sampling experiments. For a

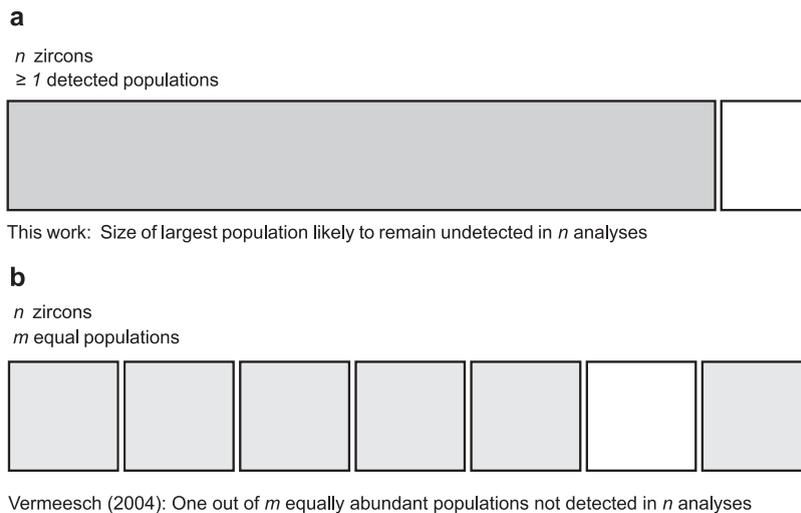


Fig. 1. Two different approaches to the definition of the detection limit for zircon populations. For both, the probability level associated with the limit (e.g., 50%, 95%) will have to be specified. (a) The detection limit is defined as the maximum size of a single population likely to remain undetected after  $n$  analyses. No assumptions are made about the number, nature, or distribution of the populations that have been detected. This approach to the problem is similar to that of Dodson et al. (1998) and is preferred in the present paper. (b) The detection limit is defined as the size of  $m$  equally distributed, equally abundant populations, one of which is likely to escape detection in  $n$  analyses. This approach was introduced by Vermeesch (2004).

single given age population, each experiment has two possible outcomes: A zircon may or may not belong to the population in question. As long as  $N \gg n$ , the relative abundance of the  $i$ th population will remain constant throughout the sampling process, and is given by  $X_i = N_i/N$ . The probability that  $n_i$  zircons selected for analysis belong to the  $i$ th age population ( $p_i$ ) is then given by the classical binomial probability formula (e.g., Zwillinger, 2003):

$$p_i = \frac{n!}{n_i!(n - n_i)!} X_i^{n_i} (1 - X_i)^{(n - n_i)}, \quad (1)$$

where  $n$  is the total number of zircons analysed, and  $n_i$  is the number among these belonging to the  $i$ th population. The probability that the zircon does *not* belong to this age population ( $q_i$ ) is  $q_i = 1 - p_i$ .

### 2.3. Detection limits: the chance of finding vs. overlooking an age population

A given age population may be said to have been detected in a set of  $n$  analyses when at least one zircon belonging to the population has been detected (i.e.,  $n_i \geq 1$ ). In the opposite case, no zircons belonging to the population are recorded ( $n_i = 0$ ). In practical work, it may be desirable to consider successfully detected populations for which  $n_i$  exceeds a certain threshold value (e.g.,  $n_{i,\min} = 3$ ; Pell et al., 1997), which will lead to a higher detection limit than the limiting case considered here. For  $n_i = 0$ , Eq. (1) reduces to:

$$p_{n_i=0} = (1 - X_i)^n \quad (2a)$$

which is the probability of overlooking an age population with an abundance  $X_i$  in the sediment (Dodson et al., 1988). This, or the complementary probability of finding (at least) one zircon belonging to the  $i$ th age population:

$$p_{n_i>0} = q_{n_i=0} = 1 - (1 - X_i)^n, \quad (2b)$$

defines the detection limit at a given level of probability for an age population. The limiting relative population size ( $X_L$ ) is thus dependent both on the number of grains analysed and on the probability level assigned to the detection limit. The abundance of an age population at the detection limit of probability  $p_L$  is obtained by solving Eq. (2b):

$$X_L = 1 - (1 - p_L)^{1/n}, \quad (3)$$

which is plotted as a function of  $n$  in Fig. 2 for probability levels of 0.5 and 0.95.

At  $p_L = 0.5$ , there is an equal probability of finding, and failing to find, at least one zircon belonging to the  $i$ th age population, whereas at  $p_L = 0.95$ , there is a 95% chance of finding at least one zircon belonging to this population. The  $p_L = 0.95$  limit may be regarded as a fairly safe indicator that a population of the corresponding true abundance will be detected in  $n$  analyses, and is thus a robust detection limit for the method (Sircombe, 2000). On the other hand, the  $p_L = 0.5$  line in Fig. 2 marks an upper abundance limit for populations that are more probably overlooked than observed in  $n$  analyses. This is further illustrated by Fig. 3, which shows the rate of failure to detect populations of 2%, 10%, 20%, and 40% abundance in 2000 Monte Carlo trials generating synthetic data sets with  $n$  ranging from 5 to 1000 (Appendix A, a). For example, an analytical study based on less than ca. 35 randomly selected zircons from each sampled unit is likely to fail to detect a population of 2% abundance in the sediment in more than 50% of the cases. On the other hand, a 10% population is slightly above the  $p_L = 0.95$  limit at  $n = 35$ , and will therefore escape detection in less than 5% of the cases.

For some applications of detrital zircon data, it is quite important to predict what may have been overlooked in a study (see Section 4, below). Therefore, the  $p_L = 0.5$  and  $p_L = 0.95$  limits should always be quoted.

### 2.4. Representativity of detrital zircon data: some theoretical limits

A set of  $n$  analysed detrital zircons may be said to be quantitatively representative of the host sediment if all the population abundances observed exactly match the corresponding abundance in the sediment. If the observed abundance of the  $i$ th population is given by  $x_i = n_i/n$ , then  $x_i = X_i$  must be true for all populations in the sediment. The probability that the observed relative abundance of the  $i$ th population and its abundance in the sediment are exactly equal can be estimated from Eq. (1) by assuming that  $n_i = nX_i$ , which is equal to the expected mean value of  $n_i$  in a series of repeated experiments (Appendix A, a). The resulting probabilities are shown in Fig. 4 as a function of  $n$  and  $X_i$ . The probability of *exactly*

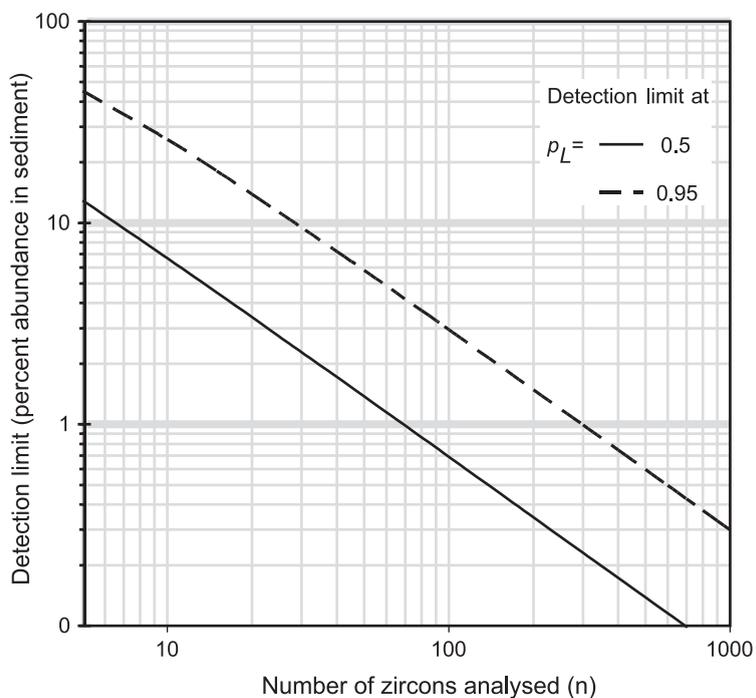


Fig. 2. Theoretical detection limits for zircon populations in data sets of  $n$  analyses, derived from the binomial formula, at probability levels  $p_L=0.5$  and  $0.95$ .

matching the predicted number of zircons is modestly high at best, reaching values of ca. 0.3 for small populations in large data sets. For  $X_i < 0.5$ , it decreases

with increasing  $n$  and  $X_i$ . This result may seem counterintuitive. However, for  $0 < X_i < 0.5$ , the binomial probability curves broaden with increasing  $n$  and  $X_i$ ,

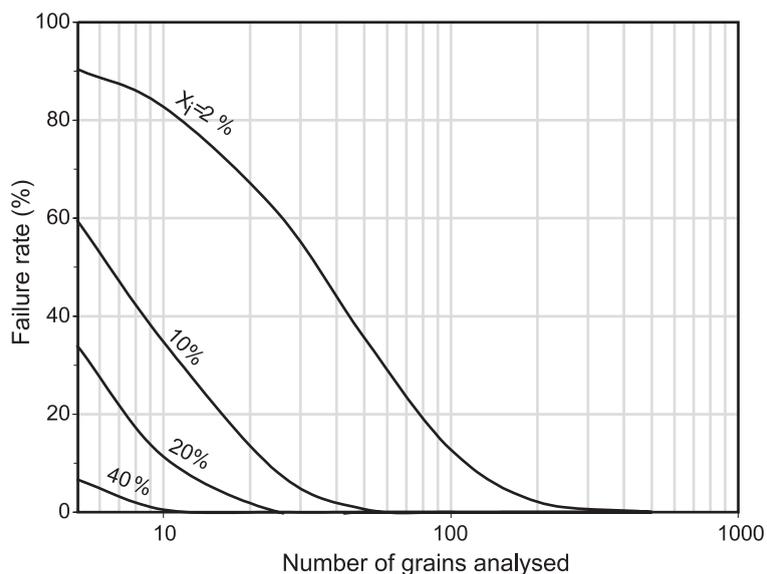


Fig. 3. The rate of failure to detect populations of given, true abundances as a function of  $X_i$  and  $n$ . The curves are based on simulated data sets with  $n$  from 5 to 1000, each involving 2000 Monte Carlo trials.

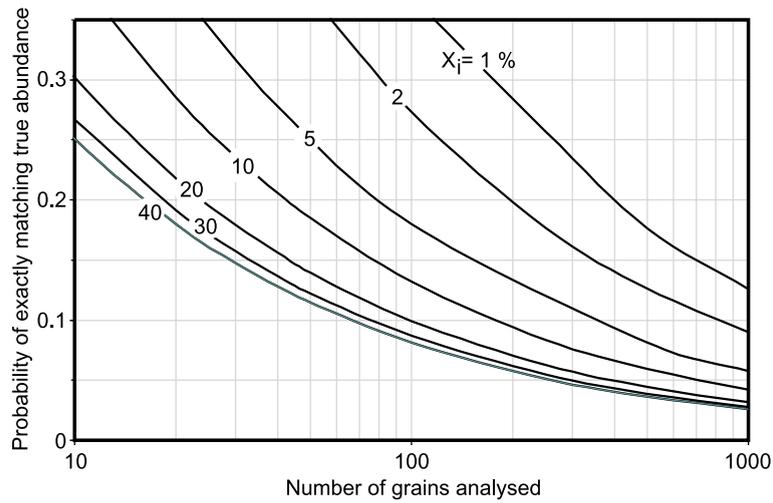


Fig. 4. The probability of a sampling experiment *exactly* matching the true population abundance (i.e.,  $n_i=nX_i$ ), calculated from the binomial formula as a function of the number of analyses and the true relative abundance of the populations.

so that the probability of *nearly* matching the true population abundance (i.e., within one or two standard deviations of the true value) will increase with increasing  $X_i$  and  $n$ .

The expected  $2\sigma$  relative error in the abundance of a zircon population in a sediment estimated by assuming that  $x_i=X_i$  can be derived from the variance of the binomial distribution, and is given

by  $2\sigma_i = 2\sqrt{nX_i(1-X_i)}$  (Appendix A, a). The expected uncertainty is illustrated as a function of  $n$  and  $X_i$  in Fig. 5. A population with an abundance of 30% would have a relative error on the 95% confidence limit of 100% or more if the total number of grains analysed is 10 or less. The relative uncertainty decreases with an increasing number of grains analysed, but even a population as abundant as

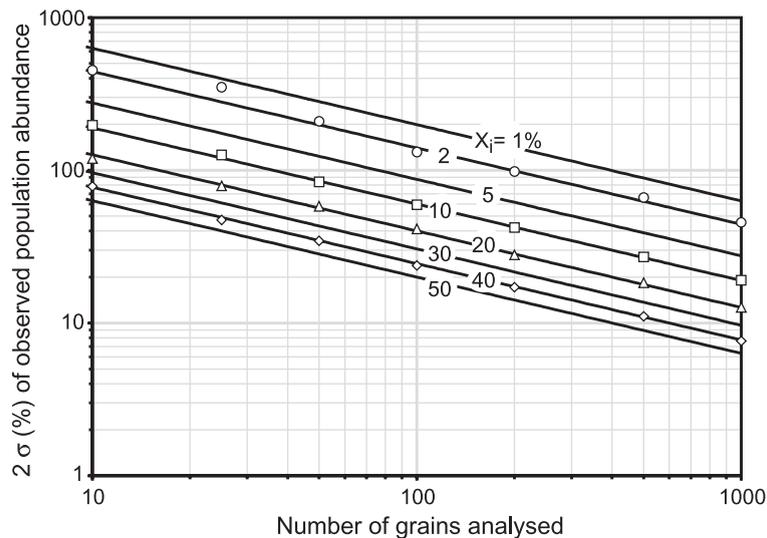


Fig. 5. The expected relative error ( $2\sigma$ ) of  $x_i$  as a function of true population abundance and the number of grains analysed. Each point represents the observed errors (two relative standard deviations) observed in synthetic data sets generated by Monte Carlo simulation (2000 trials) at  $X_i=2\%$ , 10%, 20%, and 40%.

50% will still have a  $\pm 20\%$  error at  $n=100$ . At the other extreme, a population with an abundance in the sediment less than 5% will only give an abundance estimate with an uncertainty less than 100% if more than 100 grains are analysed.

The most likely value of  $n_i$  to be observed in a single set of analyses is given by the maximum on the relevant binomial probability distribution. Unlike the Gaussian distribution, binomial distribution curves are, as a rule, asymmetrical, which means that the most frequent value of  $n_i$  observed in a series of repeated sampling experiments is not identical to the observed mean. This is reflected by a nonzero coefficient of skewness (Appendix A, a). Any zircon population with an abundance  $X_i < 0.5$  will have a positive skewness, which means that the most probable value of  $n_i$  is less than  $nX_i$ , even when randomness is rigorously observed throughout the processes. This tendency is especially pronounced for low  $X_i$ , as illustrated by the elevated failure rate in Monte Carlo simulations of a low-frequency population shown in Fig. 3. For large values of  $n$ , the shape of the binomial distribution approaches that of a normal distribution, but only very slowly for low  $X_i$  values (Fig. 6). Populations with an abundance of 10% or less will therefore tend to be underestimated, even when randomness is strictly observed throughout sample preparation and analysis.

The tendency for small populations to be underestimated in small data sets is further illustrated in Fig. 7, which shows that the most frequently returned  $n_i$  values in a series of Monte Carlo simulations (the maximum on the curves) are systematically lower than the corresponding average value (shown by vertical lines).

### 3. Numerical simulation of zircon populations

The degree to which the probability density distribution of detrital zircon ages observed in a set of  $n$  analysed zircons reflects that of the host sediment can be evaluated from synthetic data sets drawn from an infinite pool of zircons with a known distribution of ages and errors. This approach has the advantage that the overall age distribution of detrital zircons in the sediment (i.e., the pool of ages) is known, so that the relationship between the “true” age distribution and the distribution observed for the  $n$  randomly selected zircons can be evaluated as a function of  $n$ . The algorithm used to generate the data sets is briefly described in Appendix A, b.

The numerical simulations may be used to answer two questions: (1) What is the *qualitative* relationship between “true” and observed age probability density

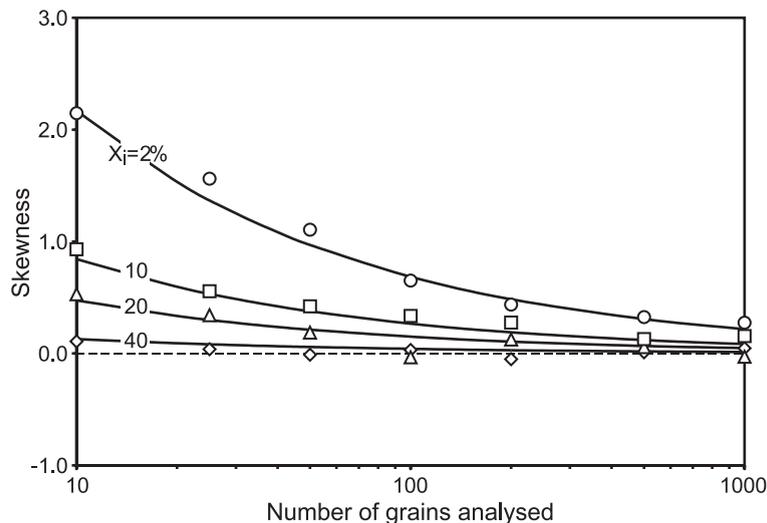


Fig. 6. Coefficient of skewness for binomially distributed  $n_i$  values as a function of true abundance (in percent) and number of grains analysed (lines). The points represent skewnesses observed in synthetic data sets generated by Monte Carlo simulation.

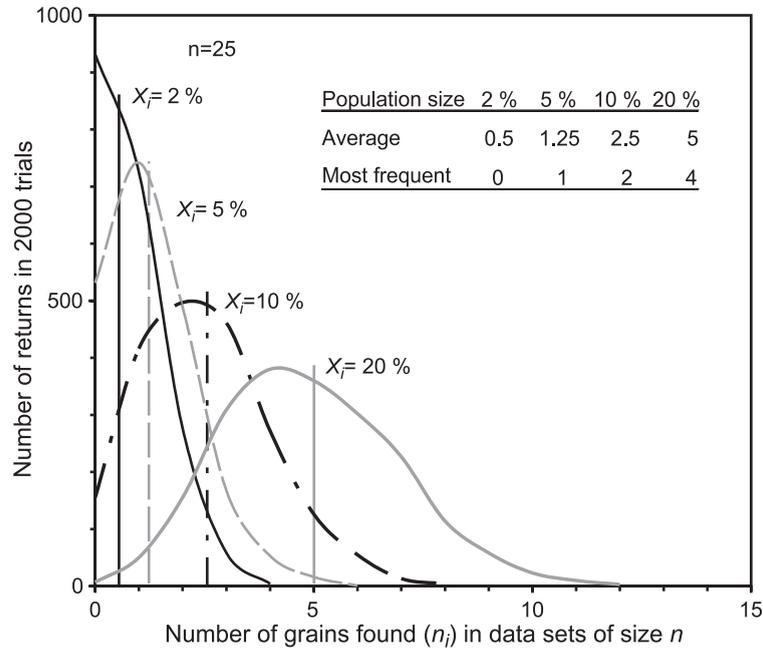


Fig. 7. Curves showing the distributions of  $n_i$  observed in 2000 Monte Carlo simulations of 2%, 5%, 10%, and 20% populations for  $n=25$ . The vertical bars represent the mean values of the respective population sizes observed after the 2000 trials. It should be noted that for all four populations, the most frequently returned number is less than the respective mean value, demonstrating the tendency for the sizes of minor populations to be underestimated in small data sets.

patterns (i.e., when will the observed ages give an unbiased representation of the age distribution pattern in the sediment, without spurious probability density maxima or minima)? (2) Is *quantitative* representation of the detrital zircon age distribution in a sediment possible for reasonable values of  $n$  (i.e., will observed age population abundances ever reflect those of the sediment)?

Most detrital zircon studies published up to 2003 base their conclusions on 20–80 analyses from each sample (e.g., Pell et al., 1997, 2001; Sircombe, 2000; Bingen et al., 2001, 2003). Most of the simulation experiments were therefore done for  $n=100$ ; this number was chosen to create statistics that are not easily surpassed by natural data. To evaluate the effect of increasing the number of analyses beyond the  $n=117$  limit of Vermeesch (2004), some experiments were made at  $n=120$ .

### 3.1. Uniform age distribution

As a first simple case, consider a sediment with a uniform distribution of zircons over a certain age

interval (Fig. 8a). Random sets of 10, 25, 50, and 100 zircon ages have been drawn from a pool consisting of zircons evenly spaced at 10 Ma age intervals, with  $1\sigma=10$  Ma, each set including the preceding smaller sets (Fig. 8b). At  $n=10$ , zircons scatter over the entire age range, without any apparent pattern. At  $n=25$ , there are clusters of zircons whose errors overlap between 1150 and 1300 Ma and between 1550 and 1750 Ma. There are prominent gaps in the distribution at 1100, 1400, 1500, and 1800 Ma (arrows). At  $n=50$ , the clusters become less obvious, although the one at 1150–1300 Ma is still recognizable as such. However, the frequency minima are still present, and persist also at  $n=100$ . At  $n=100$ , the probability density is still very irregular, suggesting that  $n \geq 100$  is required to give a reasonable representation (i.e., one without spurious peaks or valleys) of the age density distribution of the zircons in the sediment, for the distribution of ages and errors considered in this experiment.

Each of the synthetic data sets in Fig. 8b was examined for statistically significant groups of ages with the TuffZirc algorithm of Ludwig and Mundil

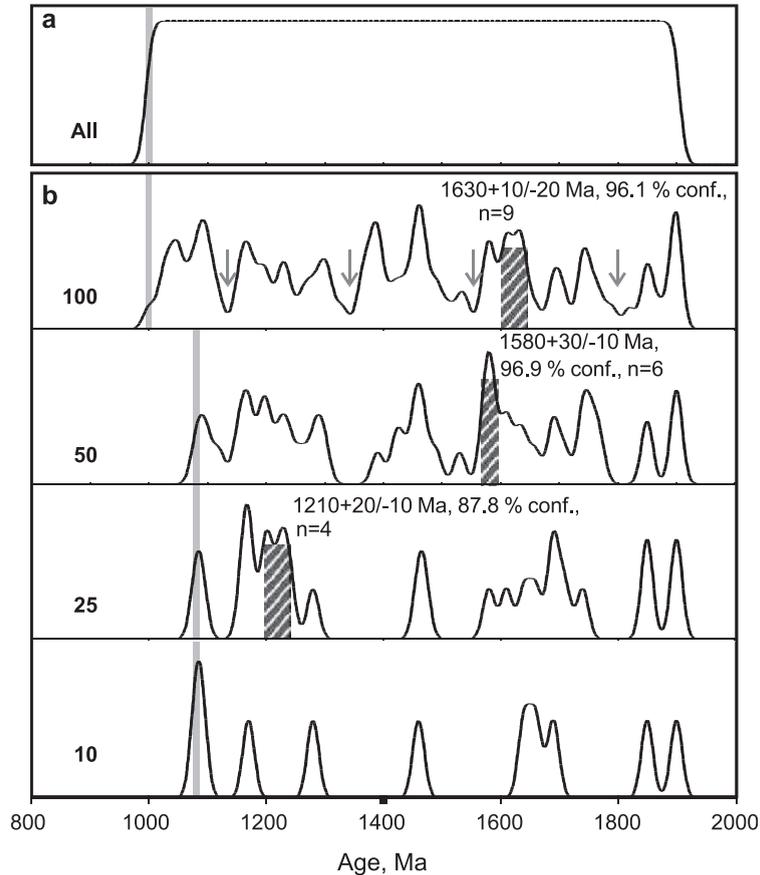


Fig. 8. Synthetic data sets generated from an infinite pool with a discrete, uniform distribution of ages. (a) Age probability density distribution of the data pool, consisting of zircons with 10 Ma age spacing between 1000 and 1900 Ma, each with  $1\sigma=10$  Ma. (b) Age probability density of synthetic data sets generated by random selection from the pool in *a* with  $n=10, 25, 50,$  and  $100$ . The shaded rectangles are clusters of ages recognized as statistically significant by the algorithm of Ludwig and Mundil (2002), at ages and errors shown.

(2002), using Isoplot 3.00 (Ludwig, 2003). Such groups were recognized at  $n=25$  (four zircons;  $1210\pm 20/-10$  Ma),  $n=50$  (six zircons;  $1580\pm 30/-10$  Ma), and  $n=100$  (nine zircons;  $1630\pm 10/-20$  Ma). These apparently significant age clusters result from an insufficient sample size, even at  $n=100$ , and could be extremely misleading if they were to occur in a natural case.

To evaluate the representativity of a “worst-case” example in the sense of Vermeesch (2004), 10 sets of 120 zircons each were drawn from a uniform age pool containing zircons at 10 Ma age intervals from 1000 to 2000 Ma, and classified in 20 evenly spaced, 50 Ma age populations (Table 1). Age uncertainties were not considered in this experiment. All 20 populations were detected in 9 of 10

experiments; an average population size of  $n_i=6$  is given by the design of the experiment, but the spread of values around this mean, expressed by one standard deviation (S.D.), is considerable. All but two of the experiments gave population sizes with a positive skewness, again illustrating the tendency of the abundance of small populations to be underestimated, even when randomness is strictly maintained.

### 3.2. Nonuniform age distribution

The simple age distribution pattern in Fig. 8a is unlikely to occur in a natural sediment. In a more realistic scenario, a sediment may contain several age populations, each of which has a nonuniform age

Table 1  
Observed  $n_i$  drawn from a uniform pool with  $X_i=5\%$

Population number	Age range (Ma)		Experiment number									
			1	2	3	4	5	6	7	8	9	10
			Observed population size ( $n_i$ )									
1	1000	1050	9	5	7	5	6	8	3	6	7	9
2	1050	1100	3	5	11	9	5	3	5	7	5	7
3	1100	1150	4	12	7	7	5	7	10	6	6	8
4	1150	1200	9	4	10	4	5	10	6	7	9	9
5	1200	1250	5	7	8	2	3	2	3	2	4	2
6	1250	1300	4	7	9	8	8	6	8	4	3	8
7	1300	1350	8	6	4	5	5	7	8	3	6	3
8	1350	1400	9	2	6	7	8	3	7	8	9	10
9	1400	1450	4	8	3	7	3	2	6	5	1	6
10	1450	1500	5	4	10	4	9	6	11	6	10	4
11	1500	1550	8	8	8	5	4	8	5	6	6	7
12	1550	1600	4	6	0	7	5	6	6	5	5	7
13	1600	1650	6	10	7	9	7	11	5	7	3	5
14	1650	1700	4	8	5	2	7	6	5	6	7	4
15	1700	1750	11	3	4	6	7	5	4	5	5	5
16	1750	1800	1	3	4	2	5	6	8	5	10	7
17	1800	1850	5	7	4	8	5	6	6	9	6	4
18	1850	1900	7	8	4	5	10	6	5	11	8	7
19	1900	1950	10	2	4	9	6	7	6	7	7	4
20	1950	2000	4	5	5	9	7	5	3	5	3	4
		$n$	120	120	120	120	120	120	120	120	120	120
		Average	6	6	6	6	6	6	6	6	6	6
		1 S.D.	2.7	2.6	2.8	2.4	1.9	2.3	2.2	2.0	2.5	2.2
		Skewness	0.25	0.36	0.02	-0.4	0.38	0.11	0.68	0.44	-0.1	0.03

distribution. The sediment in Fig. 9a contains three zircon populations of different abundance (Fig. 9b), each of which has normally distributed, random ages centered at 1700 Ma ( $\sigma=100$  Ma,  $X_1=90\%$ ), 1400 Ma ( $\sigma=80$  Ma,  $X_2=9\%$ ), and 1100 Ma ( $\sigma=30$  Ma,  $X_3=1\%$ ). The pool is based on a total of 2000 zircon ages, each of which is assigned a random age uncertainty between 5 and 20 Ma. The overall distribution pattern has a broad and nearly symmetrical maximum at 1700 Ma, with one tail extending into the early Proterozoic; on the low-age side, its tail overlaps with the 1400 Ma population, obscuring the separation between the two. The 1100 Ma population, on the other hand, is well separated from the older populations. The youngest zircon in the pool has an age of  $1032 \pm 11$  Ma. The age probability density patterns obtained in 10 sets of 100 randomly drawn zircons from this pool are shown in Fig. 9c, where they have been arranged in a sequence from *i* to *x* according to the age of the youngest zircon.

All the synthetic data sets are, as expected, dominated by the old 90% population. The accumulated probability density plots in this age interval range from approximately symmetric (patterns *iii* and *vi*) through quite jagged and irregular (*ii*, *v*, and *vii*) to apparently bimodal (*x*). Several of the patterns show subsidiary maxima or prominent shoulders at 1800 Ma (*i*, *ii*, *v*, *vii*, and *viii*). The 1400 Ma population is present in all 10 sets of data, but only patterns *viii* and *ix* show clear separation between the two; in *iii*, *v*, *vi*, and *ix*, the overlap area between the two populations shows higher probability densities than the 1400 Ma population itself. At  $n=100$ , the  $p=0.95$  detection limit is at ca. 3%, and populations less abundant than 0.7% in the sediment are more likely to be overlooked than detected. The young 1% population was detected in 5 of 10 synthetic data sets, and the youngest member of this population in one of the sets (*x*).

The synthetic data sets in Figs. 6 and 7 show numerous, spurious age probability density peaks.

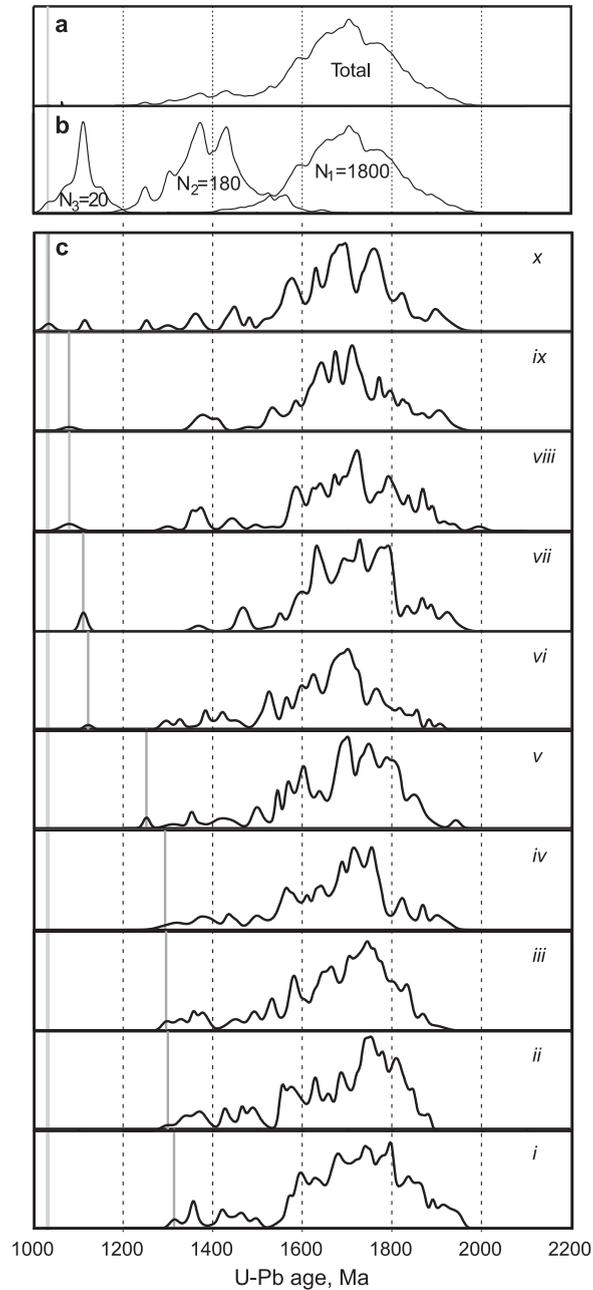


Fig. 9. Synthetic data set generated from a pool with a nonuniform distribution of ages. (a) Age probability density distribution of an infinite pool containing three populations as shown in (b), which were mixed in their true proportion. (b) Age probability density distributions of three populations, each consisting of  $N_i$  zircons. The populations were generated by randomly generated ages with normal distributions with specific mean values and standard deviations (1700 Ma, 100 Ma; 1400 Ma, 80 Ma; and 1100 Ma, 30 Ma, respectively). (c) Age probability density of synthetic data sets generated by random selection from the pool in *a* with  $n=100$ . The thin grey line continuing through all profiles represents the youngest zircon in the pool; thicker grey lines represent the youngest zircon in the synthetic data set.

These are likely to be picked up by spectrum deconvolution routines such as the [Sambridge and Compston \(1994\)](#) algorithm. This, and other quantitative methods (e.g., [Sircombe and Hazelton, 2004](#)), should therefore only be applied to sets of data when  $n$  is large enough to eliminate spurious peaks. The present example suggests that 100 randomly selected zircons will not be sufficient, even for a sediment with a comparatively simple total distribution of detrital zircon ages.

### 3.3. A seminatural case

In a final example, all analyses of detrital zircon ages from clastic (meta)sediments in S. Norway and adjacent parts of SW Sweden published up to 2003 have been combined to give a pool of zircons with a total age probability density distribution shown in [Fig. 10a](#) (see [Table 2](#) for references). This density pattern is a possible representation of a young sediment derived from the quite complex continental crust of the region (see Section 4, below). The overall pattern is constrained by 576 zircons, and shows a broad maximum between 1400 Ma and 2100 Ma, which may itself be made up by two overlapping peaks with maxima between 1500 and 1600 Ma, and in the range 1700–1900 Ma. Subsidiary populations are located (1) in the late Archaean, with indications of a peak at ca. 2.7 Ga; (2) at 900–1200 Ma; and (3) in the late Neoproterozoic to early Phanerozoic. The youngest zircon was dated at  $329 \pm 2$  Ma.

In the synthetic data sets generated from this pool, 1400–2100 Ma zircons are always dominant ([Fig. 10b](#)). However, the age probability density distribution in this age interval is quite variable, ranging from continuous without any clear grouping (patterns *ii*, *v*, and *x*) to distributions strongly indicating two distinct subpopulations (*i*, *iv*, *vii*, and *ix*). Phanerozoic zircons fail to materialize in three data sets (*i*, *ii*, and *iii*), and the youngest zircon in the pool is only recognized in one of the 10 sets (*x*). At the opposite end, pre-3.0 Ga zircons are recognized in four experiments (*ii*, *iii*, and *vi*) and late Archaean zircons in all.

To gain an impression of how well overall mass balance is maintained in the synthetic data sets, the relative abundances of zircons of three age popula-

tions (900–1200 Ma, 1400–1700 Ma, and 1700–2100 Ma) are compared (pie diagrams in [Fig. 8](#)). In the pool, these have relative abundances of 10%, 44%, and 46%, respectively. In the synthetic data sets drawn from the pool, the observed sizes of the two larger populations remain approximately equal (but see significant variations either way in patterns *v*, *vi*, and *x*), with average population sizes of  $43 \pm 5\%$  (1 standard deviation) for the 1400–1700 Ma population, and  $46 \pm 4\%$  for the 1700–2100 population. The relative size of the 900–1200 Ma population varies more widely, from 4% to 17%, with an average of  $11 \pm 4\%$ .

## 4. A case study: Southwestern Baltica

The Baltic Shield is composed of an Archaean core in the northeast, and progressively younger Proterozoic crustal domains towards the southwest (e.g., [Gaàl and Gorbatshev, 1987](#)). In the central and southwest parts of the Baltic Shield, orogenic events have been identified in the periods 1.9–1.75 Ga (Svecofennian), 1.75–1.50 Ga (Gothian), and 1.2–0.9 Ga (Sveconorwegian, i.e., Grenvillian; [Gaàl and Gorbatshev, 1987](#)). A large, roughly N–S trending belt of granitic intrusions and rhyolitic porphyries (the Trans-Scandinavian Igneous Belt, or TIB) was emplaced in the period 1.85–1.65 Ga ([Åhäll and Larson, 2000](#)), and separates the Svecofennian domain from Gothian and Sveconorwegian terranes of southwestern Baltica. There is evidence of magmatism in SW Baltica both after the end of active subduction at ca. 1.50 Ga, and in Sveconorwegian time ([Wellin, 1994](#); [Sigmond, 1998](#); [Knudsen and Andersen, 1999](#); [Andersen et al., 2002a,d](#); [Laajoki et al., 2002](#)).

[Table 2](#) summarizes age populations of detrital zircons identified in six published detrital zircon studies from the region. Five of the studies have been made by SIMS, using the Nordsim Cameca ims1270 in Stockholm ([Knudsen et al., 1997](#); [Haas et al., 1999](#); [Bingen et al., 2001, 2003](#)) or the ANU SHRIMP ([Åhäll et al., 1998](#)); the sixth ([Dahlgren and Corfu, 2001](#)) was based on TIMS-ID analysis of handpicked, single zircons. The largest data set comprises 49 zircons; the smallest, six zircons. The published U–Pb ages have been assigned to age

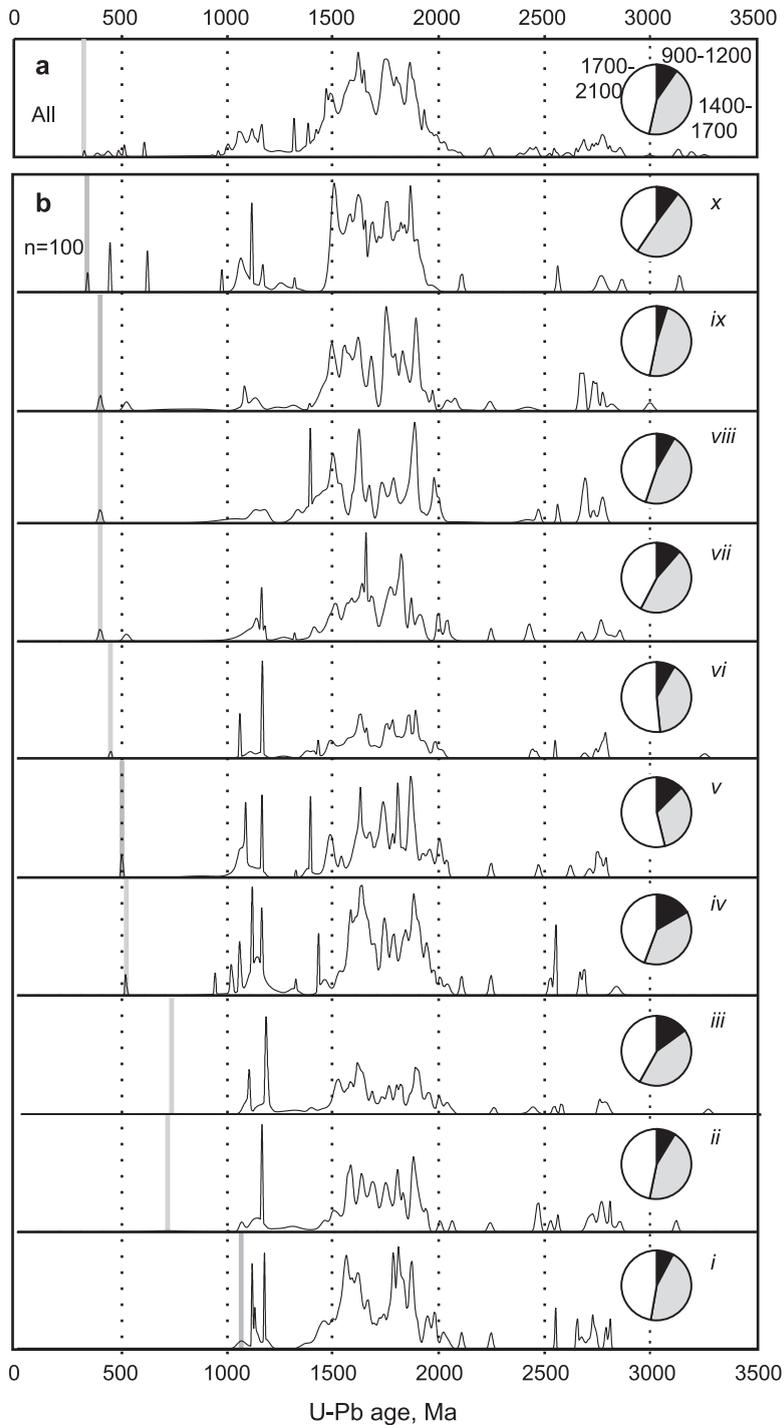


Fig. 10. Synthetic data set generated from a pool of real detrital zircon ages. (a) Age probability density distribution of an infinite pool defined by all detrital zircon ages from SW Baltica published up to 2003 (see Table 1 for references). (b) Age probability density of synthetic data sets generated by random selection from the pool in *a* with  $n=100$ . Grey, vertical lines show the ages of the youngest detrital zircon in the data sets. Pie diagrams show relative abundances of three age populations (900–1200 Ma, 1400–1700 Ma, and 1700–2100 Ma, respectively).

Table 2  
Detrital zircon age data, southwestern Baltica

Sample/complex/area	<i>n</i>	Population	<i>n<sub>i</sub></i>	Detection limits (%)		Zircon fraction size			Score (Table 3)	References
				probability level		<i>x<sub>i</sub></i> (%)	2 <i>s</i> Absolute	Relative (%)		
				<i>p<sub>L</sub></i> =0.5	<i>p<sub>L</sub></i> =0.95					
Ringerike group, Oslo Region	19	Pre Sveconorwegian	6	4	15	32	2.0	34	3	Dahlgren and Corfu (2001)
		Sveconorwegian	10			53	2.2	22	4	
		Ordovician	3			16	1.6	53	2	
Asker group, Oslo Region	16	Pre Sveconorwegian	3	4	17	19	1.6	52	2	Dahlgren and Corfu (2001)
		Sveconorwegian	4			25	1.7	43	3	
		Avalonian	3			19	1.6	52	2	
		Cambrian	3			19	1.6	52	2	
		Carboniferous	3			19	1.6	52	2	
Veme quartzite, Kongsberg sector	9	1.50–1.65 Ga	8	7	28	89	0.9	12	5	Bingen et al. (2001)
		1.65–1.85 Ga	1			11	0.9	94	2	
Modum complex, Kongsberg sector	43	<1.40 Ga	2	2	7	5	1.4	69	1	Bingen et al. (2001)
		1.40–1.50 Ga	3			7	1.7	56	2	
		1.50–1.65 Ga	7			16	2.4	35	2	
		1.65–1.85 Ga	9			21	2.7	30	3	
		1.85–2.50 Ga	18			42	3.2	18	3	
Hallingdal complex, Telemark sector	40	1.65–1.85 Ga	9	2	7	23	2.6	29	3	Bingen et al. (2001)
1.85–2.50 Ga	22			55	3.1	14	4			
>2.50 Ga	9			23	2.6	29	3			
Heddal group, Telemark sector	38	<1.40 Ga	11	2	8	29	2.8	25	3	Bingen et al. (2003)
		1.40–1.50 Ga	6			16	2.2	37	2	
		1.50–1.65 Ga	4			11	1.9	47	2	
		1.65–1.85 Ga	5			13	2.1	42	2	
		1.85–2.50 Ga	7			18	2.4	34	2	
Kalhovd group, Hardangervidda	49	>2.50 Ga	5			13	2.1	42	2	Bingen et al. (2003)
		<1.40 Ga	18	1	6	47	3.1	17	3	
		1.40–1.50 Ga	5			13	2.1	42	2	
		1.50–1.65 Ga	6			16	2.2	37	2	
		1.65–1.85 Ga	9			24	2.6	29	3	
Hettefjord group, Hardangervidda	43	1.85–2.50 Ga	7			18	2.4	34	2	Bingen et al. (2001)
		>2.50 Ga	4			11	1.9	47	2	
		1.40–1.50 Ga	1	2	7	2	1.0	99	1	
		1.50–1.65 Ga	6			14	2.3	38	2	
		1.65–1.85 Ga	3			7	1.7	56	2	
Festningsnuten group, Hardangervidda	19	1.85–2.50 Ga	18			42	3.2	18	3	Bingen et al. (2001)
		>2.50 Ga	15			35	3.1	21	3	
		1.50–1.65 Ga	3	4	15	16	1.6	53	2	
		1.65–1.85 Ga	5			26	1.9	38	3	
		1.85–2.50 Ga	10			53	2.2	22	4	
GA 521 Bandak quartzite, Telemark sector (Eidsborg fm) <sup>a</sup>	6	>2.50 Ga	1			5	1.0	97	2	Haas et al. (1999)
		1.7–1.9 Ga	2	11	39	33	1.2	58	3	
		1.4–1.6 Ga	2			33	1.2	58	3	
		1.7–1.9 Ga	2			33	1.2	58	3	
GA 548 Faurefjell fm., Rogaland	16	0.8–1.2 Ga	5	4	17	31	1.9	37	3	Haas et al. (1999)
		1.4–1.6 Ga	5			31	1.9	37	3	
		1.6–1.8 Ga	4			25	1.7	43	3	
		>2.4 Ga	2			13	1.3	66	2	

Table 2 (continued)

Sample/complex/area	<i>n</i>	Population	<i>n<sub>i</sub></i>	Detection limits (%)		Zircon fraction size			Score (Table 3)	References
				probability level		<i>x<sub>i</sub></i> (%)	2 <i>s</i> Absolute	Relative (%)		
				<i>p<sub>L</sub></i> =0.5	<i>p<sub>L</sub></i> =0.95					
GA 526 P Pebble in Seljord group conglomerate, Telemark sector (Vallar bru fm. of the Lifjell group) <sup>a</sup>	16	1.4–1.6 Ga	2	4	17	13	1.3	66	2	Haas et al. (1999)
		1.7–1.9 Ga	13			81	1.6	12	5	
		> 2.4 Ga	1			6	1.0	97	2	
GA 526 M Matrix in Seljord group conglomerate, Telemark sector (Vallar bru fm. of the Lifjell group) <sup>a</sup>	14	1.4–1.6 Ga	3	5	19	21	1.5	51	3	Haas et al. (1999)
		1.7–1.9 Ga	9			64	1.8	20	4	
		> 1.9 Ga	2			14	1.3	65	2	
GA 550 Selås metadurbidite, Bambler sector	8	1.3–1.5 Ga	2	8	31	25	1.2	61	3	Haas et al. (1999)
		1.5–1.6 Ga	2			25	1.2	61	3	
		1.7–1.9 Ga	4			50	1.4	35	4	
36.93 ITO Quartzite Bamble sector	13	1.4–1.6 Ga	7	5	21	54	1.8	26	4	Knudsen et al. (1997)
		1.6–1.8 Ga	6			46	1.8	30	3	
33.93 ITO Quartzite Bamble sector	10	1.4–1.6 Ga	3	7	26	30	1.4	48	3	Knudsen et al. (1997)
		1.6–1.8 Ga	5			50	1.6	32	4	
		1.8–2.0 Ga	2			20	1.3	63	3	
DC94-20 Quartzite (Kragero), Bamble sector	20	<1.40 Ga	2	3	14	10	1.3	67	2	Åhäll et al. (1998)
		1.40–1.50 Ga	5			25	1.9	39	3	
		1.50–1.65 Ga	4			20	1.8	45	3	
		1.65–1.85 Ga	6			30	2.0	34	3	
		1.85–2.50 Ga	2			10	1.3	67	2	
Stora Le-Marstrand formation, SW Sweden	22	1.50–1.65 Ga	19	3	13	86	1.6	8	5	Åhäll et al. (1998)
		1.65–1.85 Ga	3			14	1.6	54	2	

<sup>a</sup> Revised stratigraphy of Telemark supracrustals: Laajoki et al. (2002).

populations without considering analytical precision or concordance of individual points. For most of the data sets, a series of six standard age populations corresponding to major active periods in the Baltic Shield was chosen (<1.4 Ga, 1.4–1.5 Ga, 1.5–1.65 Ga, 1.65–1.85 Ga, 1.85–2.50 Ga, and >2.5 Ga), but for some of the smaller data sets (Knudsen et al., 1997; Haas et al., 1999, Dahlgren and Corfu, 2001), the classification preferred by the original authors has been retained.

Even in the larger data sets ( $n \geq 40$ ), the 95% confidence detection limit is as high as 7%. There is thus a real possibility that potentially significant, minor populations may have been overlooked in all of these studies. For the smaller data sets ( $n \leq 20$ ),

populations less abundant than ca. 3% are more likely to have been overlooked than detected; this is even worse for the smallest data sets, where 10% populations are below the  $p_L=0.5$  detection limit.

Only one of the population abundances calculated in Table 2 has a relative uncertainty less than 10%, and most fall between 40% and 70%. The smallest populations observed are made up by a single zircon, which amounts to 2–9% of the total data set; the relative uncertainty on the abundance of such a population exceeds 90%. Clearly, data sets with  $n < 40$  are inadequate to represent the age distribution of the sediments, and far less than the mass balance of input to the sedimentary basins. The numerical simulation experiments summarized in Figs. 8–10

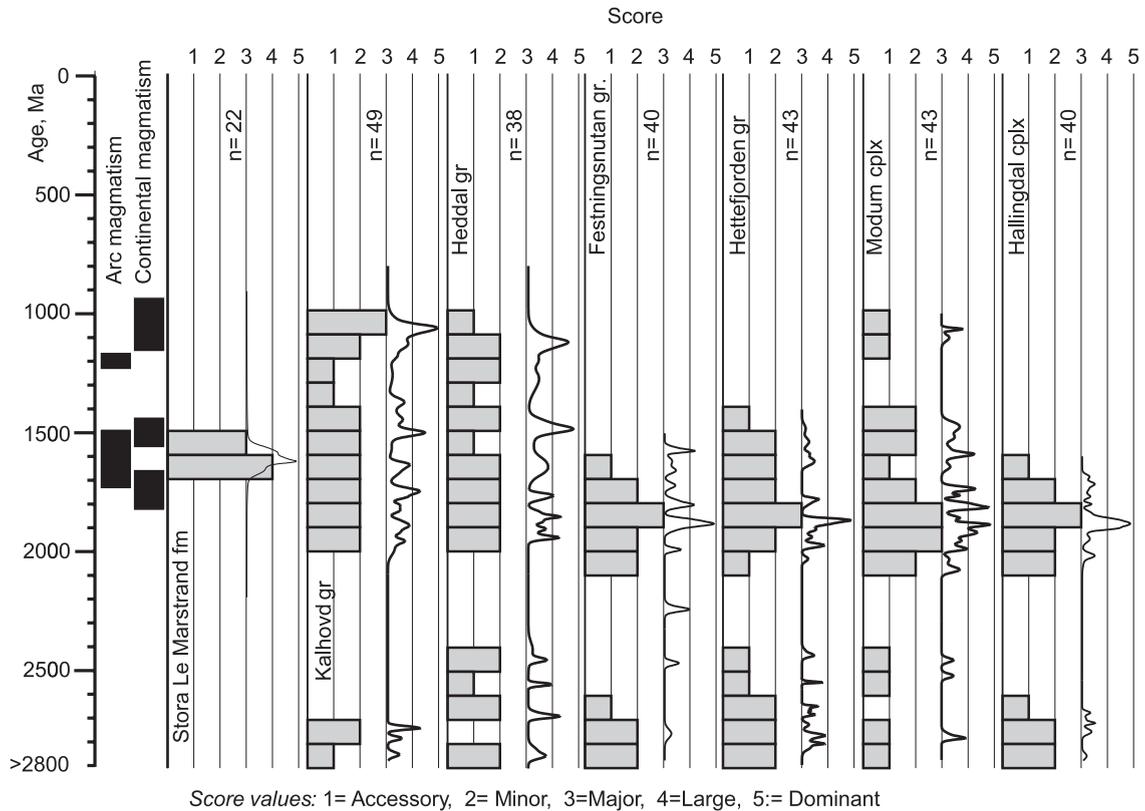


Fig. 11. Comparison of probability density profiles (lines) and qualitative abundance scores (bars) for seven samples of detrital zircons from SW Baltica (see Table 1 for sources of data, and Table 2 for definition of scores). Data have been assigned to 100 Ma bins without regard for analytical error.

indicate that age probability density diagrams constructed from these data are likely to show spurious peaks (see examples in Fig. 11).

## 5. Discussion

The numeric modelling in this study indicates that the U–Pb ages of detrital zircons are unlikely to give an accurate and precise image of the age distribution in the sedimentary rock from which the zircons have been separated, for the number of analyses commonly made in detrital zircon dating studies. As a result, the extrapolation from sample to processes in the provenance terrain may be problematic.

It should be stressed that the statistical considerations of this study must in general be applied to each individual sample, unless there is full stratigraphic

control, showing that two or more samples were collected from the same unit, or from units that are identical in age, and there is independent evidence that they have been deposited in settings that have sampled identical sources. Data from separate samples should not be combined to establish an acceptable total number of analyses. Identical provenance characteristics must be a hypothesis to be tested by data, not an assumption made to allow interpretation of the data.

### 5.1. Is detrital zircon a potential mass balance indicator?

The use of observed population abundances to evaluate the mass balance of a sedimentary basin would be an ultimate goal of detrital zircon studies. There are many factors likely to make the mass balance of zircons deviate from that of the bulk

sediment: This includes differences in modal zircon abundances among protosource rocks, fractionation of heavy from lighter minerals during transport, preferential loss of metamorphic high-U zircons due to mechanical abrasion, restricted range of grain size represented by zircon amenable to analysis, and processes during sample preparation (Pell et al., 2001; Sircombe and Stern, 2002; Andersen and Laajoki, 2003; Fedo et al., 2003).

However, even when randomness is most strictly maintained throughout sample processing and analysis, the statistical properties of zircon age distributions examined here are likely to bias the observed abundances relative to the true abundances in the sediment. Since observed zircon population sizes are binomially rather than normally distributed, a systematic, negative bias for the less abundant populations is unavoidable (Figs. 6 and 7), suggesting that the abundance of any zircon population amounting to less than ca. 10% of the zircons analysed will be systematically underestimated in data sets comprising less than ca. 100 analyses. For populations at the 1–2% abundance level, a significant skewness persists to several hundred analyses. The relative error in the population size will decrease with increasing number of analyses, but only slowly for small populations (Fig. 5), and will remain above 100% for a population size less than 5% even when 100 grains are analysed. Even a 10% population may have a ca. 30% relative error in data sets of this size (Fig. 5).

Haas et al. (1999) identified a dominant population of Paleoproterozoic zircons in their sample GA526, which is a conglomeratic quartzite from Telemark, S. Norway, that is now assigned to the Vallar bru formation of the ca. 1.15 Ga Lifjell group (Laajoki et al., 2002). 1.9–1.7 Ga zircons amounted to 81% of the zircons analysed from pebbles, and 61% of those from the matrix. Local, ca. 1.5 Ga ignimbritic rhyolites were less important as a source of zircons. Quantitative modelling of the sedimentary mass balance based on whole-rock Nd isotope data indicates that the relative proportion of 1.5 Ga volcanogenic material in this unit was 70–90% in the matrix, but lower (30–70%) in the clasts (Andersen and Laajoki, 2003). In this example, there is no obvious correspondence between the relative sizes of the analysed zircon populations and the bulk mass balance of this sediment (i.e.,  $x_i \neq X_i$ ).

### 5.2. What can be learned from small data sets?

Whereas very small data sets ( $n \leq 10$ ), such as those from the Veme quartzite (Bingen et al., 2001) and the Selås gneiss and Bandak group (Eidsborg formation) quartzite (Haas et al., 1999), clearly cannot give a good representation of zircon distribution in the sediment sampled, such data sets may still give some important information (Table 2): Only populations more abundant than ca. 26% exceed the detection limit at the  $p_L=0.95$  confidence level at  $n=10$ . A zircon population recognized in such a data set may therefore be safely assumed to be an important constituent of the sediment, at least if it is represented by more than a single grain.

### 5.3. Detrital zircons and timing of deposition

For correlation between drillholes, different sampling sites within a sedimentary basin or between different basins or tectonostratigraphic terranes, the age of deposition is a most important parameter. In the absence of fossils, or datable diagenetic minerals, the time of sediment deposition is extremely difficult to determine. Because the crystallization of the youngest detrital zircon within a sediment precedes deposition of the host sediment, the youngest detrital age obtained is commonly used as a maximum limit for the age of deposition (e.g., Knudsen et al., 1997; Bingen et al., 2001; Williams, 2001; Fedo et al., 2003).

However, this limit is, at best, a weak one, as there is no necessary connection between the timing of zircon-generating events in a protosource terrain and the age of final deposition of a sediment eroded from this terrain. For example, Andersen et al. (2002c) reported detrital zircon ages from ca. 1.0 Ga to 2.7 Ga in a Quaternary glaciofluvial sediment from central S. Norway (ca. 10,000 a); in that case, the true age of deposition postdated the maximum age by as much as ca. 1000 Ma.

Whether or not the use of the youngest zircon in a data set as a limit to the age of deposition of the host sediment is permissible depends both on the nature of the age distribution pattern and on knowledge of the geology of possible source terrains. A first requirement must be that the data set is large enough to be reasonably representative (i.e., that populations at the 1–2% level are unlikely to be overlooked) (i.e.,  $n=35-70$ ). Even when the data set is sufficiently large, the

method should only be used when the youngest potential protosource is likely to have contributed a significant fraction of the total detrital zircon budget of the sediment. This means either that the youngest zircon must belong to a significant age population in the sample, or that independent information on the geology of likely provenance terrains suggests that this should be the case. If the geology of the source terrain is such that the youngest possible protosource is also likely to yield significantly fewer zircons than other protosources, the population may be missed even when  $n \gg 100$ , in which case detrital zircon data should not be used to “date” deposition. If there is no independent information on the geology of potential protosources, great care and large numbers of analyses ( $n \gg 100$ ) should always be used. Nonrandom sampling may be used with some effect, but this will again require independent knowledge of the age and nature of potential protosources. Use of “maximum deposition ages” for correlation between wells, basins, or tectonic terranes may give very misleading results and should be avoided.

#### 5.4. Recommendations

The application of detrital zircon data to sedimentary provenance is an area where the nominal precision of an individual analysis will always be less important than the number of grains analysed, and the care by which they have been selected. Labour-intensive analytical methods such as isotope dilution (ID) TIMS is not able to produce the number of analyses required with a reasonable investment of time and manpower, nor can a slow instrumental method such as direct evaporation analysis. Use of an ion microprobe ( $\geq 10$  min/analysis) or the even faster LAM-ICPMS ( $\leq 3$  min/analysis) is not only commendable, but necessary. Published ID-TIMS studies (e.g., Dahlgren and Corfu, 2001; Barr et al., 2003; Davis, 2002) clearly do not include enough analyses to give an adequate statistical representation of the samples studied.

##### 5.4.1. Random and nonrandom sampling

The selection of which, and how many, zircons to analyse will determine the usefulness of any detrital zircon study. The present numerical modelling and statistical considerations suggest that analyses of several hundred randomly selected grains from each

sample will be necessary to give adequate control on the total age probability density pattern, the identification and relative abundance of age populations, and the presence and age of potentially limiting, minor populations. This is rarely possible, even with a LAM-ICPMS. In practical analytical work, a compromise may be both necessary and desirable: (1) A sufficient number of randomly selected grains should always be analysed from each sample, to achieve a  $p_L=0.5$  detection limit in the 1–2% range (i.e., 35–70 grains) and to suppress the tendency of skewed abundances in small populations (which may need larger numbers, depending on relative population abundances; Fig. 6). (2) In addition, a complementary set of nonrandomly selected zircons should be analysed to cover any minor, deviating population too small to be caught by the random data. Selection of zircons for nonrandom analysis causes a danger of finding only what is looked for, and requires considerable skill and patience from the operator. Careful visual observation during mounting is vital; the choice of grains may be assisted by cathodoluminescence and/or electron backscatter imaging and by electron microprobe analysis (e.g., Belousova et al., 2002). It is important to keep in mind that points from the two sets of data should be kept strictly apart, and that nonrandom analyses should never be used in a statistical analysis, or to construct a “total” age probability density diagram. Such diagrams could be grossly misleading because very small populations would be systematically overrepresented. The best way of visualizing such data would be with a combination of an age probability density pattern based on the randomly selected grains only, with superposed markers indicating the ages or age ranges of the nonrandom grains.

##### 5.4.2. An alternative way of visualizing data

Accumulated probability density plots (Sircombe, 2000, 2004; Ludwig, 2003) provide a compact graphical presentation of the analytical data from a sample, including ages, clusters, hiatuses, and the effect of analytical error. Such diagrams provide a convenient way of visualizing data obtained by random sampling of detrital zircons, and their use for this purpose is recommended. However, the results of the present numerical experiments suggest that maxima and minima observed in the probability

density plot of a sample may not be representative of the sediment from which it was sampled, at the numbers of analyses commonly made in a provenance dating study. The probability density plot may therefore be a less convenient tool for provenance analysis than is commonly assumed. For reasonably sized data sets, such diagrams should therefore not be used to support quantitative claims about the overall age distribution of the sediment, or about the provenance terrain. Use of scores derived from probability density distributions as a basis for multivariate statistical analysis (Sircombe, 2000) may also give misleading results, especially when narrow age bins are used for scoring. The use of narrow-binned, normal histograms (e.g., Vermeesch, 2004) is faced with similar problems.

When data that are not quantitatively representative of the sediment are used to derive properties of the sediment itself, or its provenance history, any potentially spurious age peak should be removed or smoothed out. However, as there is no way of distinguishing spurious peaks and minima from real ones without prior knowledge of the true distribution pattern, this may result in loss of significant information.

Rather than accumulating the excessive numbers of analyses needed to obtain an approximation to quantitative representation of the sediment ( $n \gg 100$ ), or using data filtering routines, it may be preferable to use broader age bins, and a qualitative ranking scale for population abundances. One such scale is suggested in Table 3, and has been applied to the data in Table 2 (scores column). Fig. 11 shows a comparison of age probability density diagrams and qualitative scores for 100 Ma bins for seven Precambrian metasedimentary rocks from SW Baltica. The ages have been binned without regard to the reported uncertainties on individual analyses. This is permissible as long as the bin size is large compared to the analytical error on individual grains. The overall range and main maxima

of protosource ages indicated by the probability density diagrams are reproduced by the score values, whereas the sawtooth patterns seen for some of these are not. For the number of zircons reported in each of the studies (from 20 to 49), some of the peaks and valleys suggested by the age probability density plots are likely to be artefacts caused by an insufficient number of analyses, and the 100 Ma scores may give a sufficient representation for correlation purposes. The combination of diagrams shown in Fig. 11 may be a convenient way of visualizing analytical data from an individual sample (the probability density profiles) and the provenance information that can be extracted from these data (the score columns).

### 5.5. Zircon—more than a chronometer

This discussion has been concerned with ages only, as have the majority of studies so far published. However, it should be noted that zircon is a potentially powerful petrogenetic indicator mineral, whose trace element composition and Hf isotope composition may give information on the chemistry, age, and petrogenetic history of the protosource (e.g., Belousova et al., 2002; Hoskin and Schaltegger, 2003; Kinney and Maas, 2003; Griffin et al., 2004). For example, the data in Andersen et al. (2002b, 2004) show that some of the zircons making up the dominant Mesoproterozoic to Paleoproterozoic peak in Fig. 10a differ in initial Hf isotopic composition, and can therefore be related to protosources formed in different tectonic settings. Trace element and Hf isotope data should therefore be included whenever possible.

## 6. Conclusions

Dating of detrital zircons by rapid, instrumental methods such as SIMS or LAM-ICPMS has in many respects proved its value as a method to study the provenance of clastic sediments and evolutionary processes in the provenance terrain. However, few if any studies published so far may have achieved the quantitative representation of the host sediment needed to use detrital zircon data for provenance mass balance analysis, or to apply quantitative methods such as age spectrum deconvolution or principal component analysis based on probability density scores.

Table 3  
Qualitative classification of zircon populations

Population size (%)	Classification	Score
0–5	Accessory	1
6–19	Minor	2
20–49	Major	3
50–79	Large	4
80–100	Dominant	5

Statistical analysis and numerical simulation indicate that detrital zircon geochronology is, in fact, a field where random sampling and accumulation of large quantities of data alone are unlikely to lead to an ultimate solution. Accepting that quantitative representation of the provenance history and mass balance of the host sediment by detrital zircons are unlikely to be achieved in practical work, the recommended approach would be a skilled combination of analysis of a reasonable number of randomly selected grains (at least 35–70), and of nonrandom grains thought to represent minor and/or potentially important components in the detrital zircon population. It should, however, be realized that the two types of data will answer different questions, and that they should always be analysed and interpreted separately.

The use of the youngest detrital zircon in a sediment as a limit for the age of deposition is questionable for both geological and statistical reasons, and this practice is therefore discouraged. In any case, such age limits should never be used to correlate between different drillholes, profiles, sedimentary basins, or tectonic terranes.

### Acknowledgements

Thanks are due to Professors W.L. Griffin and S.Y. O'Reilly for repeated invitations to the GEMOC Centre, Department of Earth and Planetary Sciences, Macquarie University, NSW, Australia, and to the Norwegian Research Council, the University of Oslo and Macquarie University for travel and laboratory support. Numerous discussions on the properties of zircon populations with the TerraneChron Squad at GEMOC, most notably Elena Belousova and Bill Griffin, are most gratefully acknowledged. Thanks to Stuart Graham for comments and linguistic support, and to Ian Williams and Keith Sircombe for helpful reviews. This is GEMOC publication no. 375. [PD]

### Appendix A. Generation of synthetic data sets

Two different types of synthetic data have been generated by Monte Carlo simulation.

(a) To examine the general relationships between  $n$ ,  $x_i$ , and  $X_i$  and their statistical errors, sets of

synthetic data were generated without regard for actual age values or uncertainties. The Visual Basic random number generator of Microsoft Excel was programmed to generate  $n$  uniformly distributed integers between 1 and 100. The random integers are binned according to predefined population limits. The experiment is repeated 2000 times. As expected, the resulting distributions closely resemble predicted binomial distributions (Fig. 10). The mean, variance, and skewness of binomial distributions are given by  $\mu_i = nX_i$  (mean),  $\sigma_i^2 = nX_i(1-X_i)$  (variance), and  $a_i = \frac{1-2X_i}{\sqrt{nX_i(1-X_i)}}$  (coefficient of skewness), respectively (e.g., Zwillinger, 2003).

(b) Synthetic age distribution patterns have been generated by first defining an infinitely large pool of single zircon ages with assigned uncertainties, and then using the random number generator to draw sets of  $n$  random datapoints, each consisting of an age and an assigned uncertainty. These are transferred to a separate spreadsheet, from which age probability density plots have been generated using Isoplot 3.00 (Ludwig, 2003). The distribution of points within the pool of analyses can be varied at will. The synthetic data set shown in Fig. 6 consists of a set of discrete, equally spaced ages with identical uncertainties. On the other hand, the data used to construct Fig. 7 were generated by combining subsets of normally distributed ages clustered at predefined ages and with predefined variances (Fig. 7). The uncertainty assigned to each datapoint is a randomly chosen number between 5 and 20 Ma, assuming uniform distribution. The pool of analyses used to generate Fig. 8 consists of all analyses used to generate Table 1 and Fig. 9, which have been compiled with their reported  $1\sigma$  errors, without further manipulation. As a result, some highly precise TIMS data have been included; when picked for synthetic data sets, these tend to stand out as spikes on a smoother background.

### References

- Åhäll, K.-I., Larson, S.Å., 2000. Growth-related 1.85–1.55 Ga magmatism in the Baltic Shield: a review addressing the tectonic characteristics of Svecofennian, TIB 1-related, and Gothian events. *GFF* 122, 193–206.
- Åhäll, K.-I., Cornell, D.H., Armstrong, R., 1998. Ion probe zircon dating of metasedimentary units across the Skagerrak: new

- constraints for early Mesoproterozoic growth of the Baltic Shield. *Precambrian Res.* 87, 117–134.
- Andersen, T., Laajoki, K., 2003. Provenance characteristics of Mesoproterozoic metasediments from Telemark, South Norway: a Nd-isotope mass balance model. *Precambrian Res.* 126, 95–122.
- Andersen, T., Andresen, A., Sylvester, A.G., 2002a. Timing of late- to postorogenic Sveconorwegian granitic magmatism in the Telemark and Rogaland–Vest Agder sectors, S. Norway. *Bull.-Nor. Geol. Unders.* 440, 5–18.
- Andersen, T., Griffin, W.L., Pearson, N.J., 2002b. Crustal evolution in the SW part of the Baltic Shield: the Hf isotope evidence. *J. Petrol.* 43, 1725–1747.
- Andersen, T., Griffin, W.L., Pearson, N.J., O'Reilly, S.Y., 2002c. Terrane chronology of a glaciated continent: a case study from the Kongsberg sector, South Norway, 25. Nordic Geological Winter Meeting, Reykjavik, January 9, 2002.
- Andersen, T., Sylvester, A.G., Andresen, A., 2002d. Age and petrogenesis of the Tinn granite, Telemark, south Norway, and its geochemical relation to metarhyolite of the Rjukan Group. *Bull.-Nor. Geol. Unders.* 440, 19–26.
- Andersen, T., Griffin, W.L., Jackson, S.E., Knudsen, T.-L., 2004. Mid-Proterozoic magmatic arc evolution at the southwest margin of the Baltic Shield. *Lithos* 73, 289–318.
- Barr, S.M., Davis, D.W., Kamo, S., White, C.E., 2003. Significance of U–Pb detrital zircon ages in quartzite from peri-Gondwanan terranes, New Brunswick and Nova Scotia, Canada. *Precambrian Res.* 126, 123–145.
- Belousova, E., Griffin, W.L., O'Reilly, S.Y., Fisher, N.I., 2002. Igneous zircon: trace element composition as an indicator of source rock type. *Contrib. Mineral. Petrol.* 143, 602–622.
- Bingen, B., Birkeland, A., Nordgulen, Ø., Sigmond, E.M.O., 2001. Correlation of supracrustal sequences and origin of terranes in the Sveconorwegian orogen of SW Scandinavia: SIMS data on zircon in clastic metasediments. *Precambrian Res.* 108, 293–318.
- Bingen, B., Nordgulen, Ø., Sigmond, E.M.O., Tucker, R., Mansfeld, J., Høgdahl, K., 2003. Relations between 1.19–1.12 Ga continental magmatism, sedimentation and metamorphism, Sveconorwegian province, S Norway. *Precambrian Res.* 124, 215–241.
- Dahlgren, S., Corfu, F., 2001. Northward sediment transport from the late Carboniferous Variscian Mountains: zircon evidence from the Oslo Rift, Norway. *J. Geol. Soc. (Lond.)* 158, 29–36.
- Davis, D.W., 2002. U–Pb geochronology of Archaean metasedimentary rocks in the Pontiac and Abitibi subprovinces, Quebec, constraints on timing, provenance and regional tectonics. *Precambrian Res.* 115, 97–117.
- Dodson, M.H., Compston, W., Williams, I.S., Wilson, J.F., 1988. A search for ancient detrital zircons in Zimbabwean sediments. *J. Geol. Soc. (Lond.)* 145, 977–983.
- Fedo, C.M., Sircombe, K.N., Rainbird, R.H., 2003. Detrital zircon analysis of the sedimentary record. In: Hanchar, J.M., Hoskin, P.W.O. (Eds.), *Zircon, Reviews in Mineralogy and Geochemistry*, vol. 53, pp. 277–303.
- Gaál, G., Gorbatshev, R., 1987. An outline of the Precambrian evolution of the Baltic Shield. *Precambrian Res.* 35, 15–52.
- Goode, J.W., Myrow, P., Williams, I.S., Bowring, S.A., 2002. Age and provenance of the Beardmore Group, Antarctica: constraints on Rodinia supercontinent breakup. *J. Geol.* 110, 393–406.
- Griffin, W.L., Belousova, E.A., Shee, S.R., Pearson, N.J., O'Reilly, S.Y., 2004. Archaean crustal evolution in the northern Yilgarn Craton: UPb and Hf-isotope evidence from detrital zircons. *Precambrian Res.* 131, 231–282.
- Haas, G.J.L.M. de, Andersen, T., Vestin, J., 1999. Detrital zircon geochronology: new evidence for an old model for accretion of the Southwest Baltic Shield. *J. Geol.* 107, 569–586.
- Hoskin, P.W.O., Schaltegger, U., 2003. The composition of zircon and igneous and metamorphic petrogenesis. In: Hanchar, J.M., Hoskin, P.W.O. (Eds.), *Zircon, Reviews in Mineralogy and Geochemistry*, vol. 53, pp. 27–62.
- Kinney, P.D., Maas, R., 2003. Lu–Hf and Sm–Nd isotope systems in zircon. In: Hanchar, J.M., Hoskin, P.W.O. (Eds.), *Zircon, Reviews in Mineralogy and Geochemistry*, vol. 53, pp. 327–341.
- Knudsen, T.-L., Andersen, T., 1999. Petrology and geochemistry of the Tromøy gneiss complex, south Norway, an alleged example of Proterozoic depleted lower continental crust. *J. Petrol.* 40, 909–933.
- Knudsen, T.-L., Andersen, T., Whitehouse, M.J., Vestin, J., 1997. Detrital zircon ages from southern Norway—implications for the Proterozoic evolution of the southwestern part of the Baltic Shield. *Contrib. Mineral. Petrol.* 130, 47–58.
- Laajoki, K., Corfu, F., Andersen, T., 2002. Lithostratigraphy and U–Pb geochronology of the Telemark supracrustals in the Bandak-Sauland area, Telemark, South Norway. *Nor. Geol. Tidsskr.* 82, 119–138.
- Ludwig, K.R., 2003. User's Manual for Isoplot 3.00. A Geochronological Toolkit for Microsoft Excel. Berkeley Geochronology Center, Special Publication No. 4a, Berkeley, CA.
- Ludwig, K.R., Mundil, R., 2002. Extracting reliable U–Pb ages and errors from complex populations of zircons from Phanerozoic tuffs. Goldschmidt Conference Abstracts 2002. *Geochim. Cosmochim. Acta* 66 (15A), A463.
- Pell, S.D., Williams, I.S., Chivas, A.R., 1997. The use of protolith zircon-age fingerprints in determining the protosource areas for some Australian dune sands. *Sediment. Geol.* 109, 233–260.
- Pell, S.D., Chivas, A.R., Williams, I.S., 2001. The Mallee Dune-field: development and sand provenance. *J. Arid Environ.* 48, 149–170.
- Sambridge, M.S., Compston, W., 1994. Mixture modelling of multi-component data sets with application to ion-probe zircon ages. *Earth Planet. Sci. Lett.* 128, 373–390.
- Sigmond, E.M.O., 1998. Geologisk kart over Norge, Berggrunnskart ODDA, M 1:250,000. Norges Geologisk Undersøkelse, Trondheim.
- Sircombe, K.N., 2000. Quantitative comparison of geochronological data using multivariate analysis: a provenance study example from Australia. *Geochim. Cosmochim. Acta* 64, 1593–1619.

- Sircombe, K.N., 2004. Age display: an EXCEL workbook to evaluate and display univariate geochronological data used binned frequency histograms and probability density distributions. *Comput. Geosci.* 30, 21–31.
- Sircombe, K.N., Hazelton, M.L., 2004. Comparison of detrital zircon age distributions by kernel functional estimation. *Sediment. Geol.* 171, 91–111.
- Sircombe, K.N., Stern, R.A., 2002. An investigation of artificial biasing in detrital zircon U–Pb geochronology due to magnetic separation in sample preparation. *Geochim. Cosmochim. Acta* 66, 2379–2397.
- van Wyck, N., Williams, I.S., 2002. Age and provenance of basement metasediments from the Kubor and Bena Bena Blocks, central Highlands, Papua New Guinea: constraints on the tectonic evolution of the northern Australian cratonic margin. *Aust. J. Earth Sci.* 49, 565–577.
- Vermeesch, P., 2004. How many grains are needed for a provenance study? *Earth Planet. Sci. Lett.* 224, 351–441.
- Wellin, E., 1994. Isotopic investigations of Proterozoic igneous rocks in south-western Sweden. *Geol. Fören. Stockh. Förh.* 116, 75–86.
- Williams, I.S., 2001. Response of detrital zircon and monazite, and their U–Pb isotopic systems, to regional metamorphism and host-rock partial melting, Cooma Complex, southeastern Australia. *Aust. J. Earth Sci.* 48, 557–580.
- Zwillinger, D. (Ed.), *CRC Standard Mathematical Tables and Formulae*. Chapman and Hall/CRC Press, Boca Raton, FL. 910 pp.