Distinguishing incorrect from inaccurate exposure ages by correlation between ^{26}Al and ^{10}Be measurements

Greg Balco [a,*]

[a] *Quaternary Research Center and Department of Earth and Space Sciences*
University of Washington, Mail Stop 351310, Seattle, WA 98195-1310 USA

Abstract

Scatter in cosmogenic-nuclide exposure age measurements from a single landform is caused by: i) measurement uncertainty in determining the actual nuclide concentration in a sample, and ii) geomorphic processes, such as erosion or cover by soil or snow, which cause the nuclide concentration not to reflect the true landform age. Measurement uncertainty can be reduced by making measurements on many samples from the same landform and averaging them; true scatter caused by geomorphic processes cannot be eliminated in this way. It is important to determine which effect is responsible for scatter in a particular data set, so that the ages that are *incorrect* due to geomorphic processes can be excluded, and those that are merely *inaccurate* due to measurement uncertainty can be averaged to yield a better age estimate for the entire landform. Measuring multiple nuclides, for example both ^{10}Be and ^{26}Al, in the same samples makes this possible. This is because geomorphic processes that affect the exposure age of different samples differently will produce a correlation between ^{10}Be and ^{26}Al exposure ages, but measurement uncertainties will not. Thus, incorrect ages can be distinguished from inaccurate ones by examining the correlation between ^{10}Be and ^{26}Al measurements.

Key words: Cosmogenic nuclides, beryllium-10, aluminum-26, exposure dating

1 Scatter in exposure-age measurements

Surface exposure dating is the practice of measuring trace isotopes of several elements that are produced by cosmic-ray bombardment of rocks and minerals exposed at the Earth's surface. These nuclides accumulate at an approximately constant rate, and the cosmic-ray flux is almost completely absorbed in the upper several meters of the Earth's surface, so the concentration of these nuclides reflects the time since the rock in question was exhumed from several meters depth and delivered to the surface. Thus, exposure-dating methods can be used to date a wide variety of geologic events which bring fresh rock to the surface. These include advances and retreats of glaciers, which quarry rock from the glacier bed and deposit it at the ice margin; earthquakes that bring fresh rock to the surface via slip along faults; volcanic eruptions which emplace lava flows and pyroclastic debris; and other processes that create landforms out of freshly exhumed rock. Cerling and Craig (1994) and Gosse and Phillips (2001) review many applications of exposure dating.

One challenge of accurate exposure dating is that exposure ages on multiple samples from a single landform are nearly always scattered. This scatter arises from two sources. First, measurement uncertainty, which typi-

---

* Corresponding author. Tel. 206-221-2579
Email address: `balcs@u.washington.edu` (Greg Balco).

cally amounts to several percent of the age, causes scatter even if all samples truly contain the same nuclide concentration. Second, geomorphic processes which affect the duration or intensity of exposure of each sample differently cause the true nuclide concentration, and therefore also the measured nuclide concentration, to differ among samples. These latter processes include episodic soil or snow cover, which could have a different shielding effect on boulders of different heights or shapes; surface erosion, which could have different effects on the nuclide inventory in rock types that were more or less susceptible to erosion; or downslope movement of boulders, which could change the exposure geometry of some boulders but not others. In addition, some cobbles or boulders on a landform might contain nuclide inventories inherited from previous periods of exposure. The important difference between these two sources of scatter is that the first one, scatter due to measurement uncertainty, can be compensated for by averaging many measurements. If the measurements of nuclide concentration are independent and normally distributed, which is for all practical purposes true given common laboratory practices and AMS techniques, then the mean of several measurements will yield a more precise age for the landform according to basic statistical arguments. The second source of scatter, true scatter caused by either inheritance or postdepositional geologic processes, cannot be dealt with by averaging, because there is no assurance that these processes lead to true nuclide concentrations that are evenly distributed about the true landform age.

In practice, the degree of scatter among exposure ages from a single landform is nearly always well in excess of that expected from measurement uncertainty alone (e.g., Putkonen and Swanson, 2003). Thus, one of the main challenges in exposure dating is to separate measurements that yield an *incorrect* exposure age due to geomorphic processes from those that yield an *inaccurate* exposure age due to measurement uncertainty. If the incorrect measurements can be excluded, then the inaccurate measurements can be averaged to yield a more precise estimate of the landform age.

2 Existing strategies for explaining and reducing scatter

Nearly every exposure-dating study that includes multiple exposure-ages from a single landform has discussed this issue at some length. Some existing strategies for dealing with the problem follow. The citations in this section are intended to provide a few illustrative examples, and are far from comprehensive.

Statistical measures of scatter. Several authors have compared how the observed scatter in their exposure ages compares with the amount of scatter expected from measurement uncertainty. Some authors have simply asserted *ad hoc* that the degrees of scatter are similar (e.g., Balco et al., 2002; Schaefer et al., 2002; Rinterknecht et al., 2004). Others have employed statistical measures, most commonly the reduced chi-squared statistic, which has a value near 1 if the observed and expected scatter are similar, and a higher value if the observed scatter exceeds the expected scatter. Reduced chi-squared values for sets of exposure ages from alpine glacial moraine boulders, for example, are generally much higher than 1, suggesting that most of the scatter in the ages is true scatter due to geologic processes. Often, reduced chi-squared values near 1 have been used to argue that the scatter can be entirely explained by measurement uncertainty, which means that true geologic scatter is unimportant by comparison, and the individual ages can therefore be averaged to obtain a more precise landform age (e.g., Balco and Schaefer, 2006; Barrows et al., 2002).

Statistical justification for discarding outliers. Some authors have used a statstical criterion (commonly Chauvenet's criterion; e.g., Rinterknecht et al., 2006) to detect outliers whose distance from a cluster of other data makes it very unlikely that they belong to the same distribution. As with any use of similar statistical criteria, the key weakness of this approach is that one must assume in advance that the measurements belong to a particular type of probability distribution.

Subsampling based on geomorphic arguments. Many authors have used geomorphic observations to argue that exposure ages on glacial moraines are more likely to have a young bias due to moraine erosion and boulder weathering, and therefore the true age of the moraine is best approximated by the oldest sample in a data set. Examples of this include Schaefer et al. (2006) and Phillips et al. (1996). Putkonen and Swanson (2003) considered this idea at length and provided guidelines for how to apply the idea to alpine glacial moraines. On the other hand, Benson et al. (2004) argued that nuclide inheritance was more important than postdepositional erosion, and suggested that the youngest samples in their data sets best represented the true moraine ages. Finally, Benson et al. (2007) considered exposure ages from both recent and late-glacial moraines at the same site with detailed observations of the sources, transport mechanisms, and condition of moraine boulders, concluded that both inheritance and postdepositional processes were important, and took the best estimate of the true moraine age to be near the mean of the exposure ages.

Geomorphic models. Selecting either the oldest or the youngest samples on the basis of geomorphic process arguments are special cases of the idea that a model of the processes that deliver boulders to a moraine and subsequently act to move, bury, exhume, or weather them can be used to predict the distribution of exposure ages on a moraine. One then can compare a measured age distribution to model distributions to infer the true depositional age that best explains the observations. Putkonen and Swanson (2003), and Hallet and Putkonen (1994) have considered this idea in more detail.

Bald-faced assertion. Some authors (e.g., Balco et al., 2002; Clark et al., 2003) have simply asserted that some of their measurements are outliers, excluded them, and averaged the rest.

3 Correlation between ages derived from multiple nuclides

The new strategy for dealing with scatter in exposure ages described here exploits the fact that multiple cosmogenic nuclides are produced simultaneously in the same mineral grains. The most commonly used pair of nuclides is that of ^{26}Al and ^{10}Be, both of which are produced in quartz. The production rate of ^{26}Al is 6 times that of ^{10}Be, but they are produced by similar reactions, and their production rates vary proportionately with shielding, geometric effects, or temporal changes in the cosmic-ray flux. However, measurements of ^{26}Al and ^{10}Be in a particular sample are independent. In common laboratory practice, the only measured quantity which is used in determining both the ^{26}Al and ^{10}Be concentrations of the sample is the sample weight. In comparison to the other uncertainties involved in both measurements, this measurement is for all practical purposes exact. Thus, if we measure both ^{26}Al and ^{10}Be in a set of samples where all samples truly have identical exposure histories and therefore identical nuclide concentrations, exposure ages derived from ^{26}Al and ^{10}Be measurements will be uncorrelated. On the other hand, any process that causes variation in the true nuclide concentration among samples, that is, any of the geologic processes of sample cover, erosion, or disturbance, will affect the exposure ages derived from ^{26}Al and ^{10}Be measurements together.

In a data set where the scatter is due to measurement uncertainty alone, ^{26}Al and ^{10}Be ages will be uncorrelated. If the scatter is due to geologic processes, they will be correlated.

4 Proposed pruning and averaging scheme

This observation, that if ^{26}Al and ^{10}Be ages from a single landform are uncorrelated, their scatter is due to measurement uncertainty and they can justifiably be averaged to yield a more precise age for the landform, suggests the following procedure for evaluating exposure age data sets:

(1) Determine whether the ^{26}Al and ^{10}Be measurements are correlated. A straightforward statistic for this purpose is the p-value of the linear correlation coefficient (e.g., P_c in Bevington and Robinson, 1992). P_c is the probability that an observed linear correlation between two variables could arise from uncorrelated data by chance. For example, P_c =

0.5 indicates that the observed correlation would be expected to occur 50% of the time, even if no correlation existed. If P_c is high, the scatter in the data set is most likely the result of measurement uncertainty, and all the measurements can justifiably be averaged to yield a more precise age for the landform.

(2) If the ^{26}Al and ^{10}Be measurements are correlated, remove the measurements that contribute most to the correlation. An unacceptable (that is, low) value of P_c suggests that some of the samples have nuclide concentrations that do not reflect the true exposure age of the landform. The logical way to remove these incorrect measurements is to prune the data set by discarding the measurement whose removal yields the greatest reduction in correlation, that is, the largest increase in P_c.

Repeating steps 1 and 2 until the data set has an acceptable P_c (given that there are enough samples in the data set) should eventually yield an uncorrelated data set. The examples below suggest that $P_c > 0.5$ is a reasonable threshold of 'acceptability.' This subset of the measurements can then justifiably be averaged to yield a more precise age for the landform.

This suggested procedure appears to have several advantages over other approaches to scattered exposure age data. It does not depend on any presuppositions about either the statistical distribution that exposure age data ought to follow or the geomorphic processes that are active at the site. It only allows a landform age to be computed if a significant fraction of the data set consists of samples that represent the true landform age – that is, if only some of the measurements are incorrect due to geomorphic processes. If this is not true, and none of the samples reflect the true landform age, then the procedure would never yield an uncorrelated subset of the data. Furthermore, it identifies which of the samples appear to have incorrect rather than inaccurate exposure ages, which provides insight into the geomorphic processes that could have disturbed these samples. This suggested procedure is only one of many possible ways of exploiting the basic idea that inheritance and postdepositional disturbance lead to correlated ^{26}Al and ^{10}Be measurements, and, in fact, it is possible to create synthetic data sets that would cause it to fail (for example, samples whose exposure history included long periods of burial could produce a data set where different subsets of the data would show positive and negative correlations, thus confusing the algorithm). However, it appears to accomplish the goal of distinguishing incorrect from inaccurate measurements in typical exposure-age data sets.

5 Examples

One disadvantage of this approach is that very few exposure-dating studies have measured multiple nuclides in the same samples. As Be and Al are extracted simultaneously in common laboratory practice, this is mainly due to the increased expense of making two isotope ratio measurements rather than one.

Schaefer et al. (2002) measured ^{26}Al and ^{10}Be in five samples from the Litang moraine in Tibet (Figure 1). ^{26}Al and ^{10}Be exposure ages are uncorrelated (P_c = 0.99). Likewise, Gosse et al. (1995) measured ^{10}Be and ^{26}Al in 6 boulders from the Titcomb Lakes moraine in Wyoming, and these ^{26}Al and ^{10}Be ages are also uncorrelated (Figure 2; P_c = 0.77). This indicates that the scatter in these sets of ages is most likely the result of measurement uncertainty alone, and the authors are entitled to take the average of the measurements as a better estimate of the moraine age.

Balco et al. (2002) measured ^{10}Be and ^{26}Al in 10 boulders from the Buzzards Bay moraine, and 12 boulders from the Martha's Vineyard moraine.

The data from the Buzzards Bay moraine are tightly grouped and have a reduced chi-squared value of 2.3. These observations led the authors to argue that the scatter was fully explained by measurement error, and to take the error-weighted mean and uncertainty of the full data set as the true age of the landform. A closer look at these data (Figure 3) suggests that this was unwarranted. The ^{26}Al and ^{10}Be measurements are correlated (P_c = 0.14), suggesting that much of the scatter reflects true scatter in the nuclide concentrations due to geologic processes. Removing two samples according to the scheme above (Figure 3) makes little difference in the correlation (P_c = 0.2); however, removing a third sample significantly reduces the correlation. After four samples have been discarded, the measurements are essentially uncorrelated ($P_c >$ 0.8). This suggests that the remaining measurements are scattered only by analytical uncertainty, and that the mean of these remaining ages (19,670 ± 270) is a better estimate of the true age of the moraine than the mean of the entire data set (19,230 ± 200). This exercise also provides some diagnostic information as to the possible cause of scatter in the data set. The first three samples discarded, the ones that make the largest contribution to the correlation between ^{26}Al and ^{10}Be ares, are the youngest in the data set. This suggests that i) the scatter in the data set is the result of postdepositional disturbance rather than nuclide inheritance, and ii) these three samples suffered either more snow or soil cover, or more postdepositional erosion, than the others. This second hypothesis could be tested by additional field observations. For example, returning to the field notes for this study reveals that the four smallest boulders also have the four youngest ^{10}Be exposure ages, implicating snow or soil cover as a possible cause of scatter. Measurements of surface roughness, grain size, or rock hardness might help to evaluate whether or not these samples also experienced anomalous surface erosion.

The data from the Martha's Vineyard moraine are widely scattered, with a central grouping near 23,000 years B.P. and both old and young outliers (Figure 4). The authors originally dealt with this by simply asserting that the oldest three and the youngest two ages were incorrect, and averaging the remaining ages. The full set of ^{26}Al and ^{10}Be ages are highly correlated ($P_c < 0.01$), suggesting that most of the scatter is geologic in nature. Progressively discarding samples according to the scheme above eventually results in an uncorrelated data set ($P_c > 0.9$) after discarding 7 samples. The error-weighted mean of the remaining ages is 23,350 ± 290. In this example, the correlation analysis reveals that both inheritance and postdepositional disturbance contribute to the scatter in the data set. The first five samples discarded on the basis of their contribution to the correlation between ^{26}Al and ^{10}Be measurements are the same samples that the authors originally discarded based solely on an appeal to common sense, so in this case the correlation analysis merely confirms that the authors actually possessed some common sense. Once again, however, the identification of certain samples as having incorrect rather than inaccurate ages provides testable hypotheses about the processes responsible for disturbance of the ages, that could be evaluated by additional field observations.

6 Further suggestions.

The argument above – that incorrect exposure ages can be distinguished from merely inaccurate ones on the basis of the correlation between ^{26}Al and ^{10}Be – has one significant weakness. It would fail to identify a situation where all samples had been subjected to geologic processes that acted to reduce the true nuclide concentrations equally. In this case, one might still underestimate the true age of the landform. This could occur, for example, if all samples were composed of the same rock type and were eroding at precisely the same rate, or if all samples were covered by the same thickness of snow for the same length of time. To counter this, one might argue that very special conditions are needed for these processes to affect all samples equally. For example, erosion is likely to affect different rock types and surface orientations differently, and any significant degree of erosion should result in increased scatter and detectable correlation of ages. For example, at the Buzzards Bay moraine, all the samples had the same lithology. However, if the observed correlation between ^{26}Al and ^{10}Be measurements was the result of small amounts of erosion, then erosion must have been variable among samples. This suggests that only in situations where erosion is negligible would it even be possible to get an uncorrelated result. A moraine where erosion had been so severe that no pristine boulders remained would always yield a correlated data set.

However, the possibility that all samples were affected equally by postdepositional disturbance suggests that exposure-dating samples from a single landform should be chosen so that the variation in the effect of different geomorphic processes would be maximized. That is, one should sample boulders with different heights (which would have had different amounts of snow or sediment cover), or different rock types or slope angles (which would erode at different rates). This would ensure that, if geomorphic disturbance was important, correlation between ^{26}Al and ^{10}Be measurements would be maximized. Then, if no correlation existed, one would be more confident in averaging the data set.

7 Acknowledgements.

Balco was supported during the period of this work by a postdoctoral fellowship from the National Science Foundation Office of Polar Programs. John Gosse graciously provided ^{26}Al measurements and additional source data for the Titcomb Basin moraines.

References

Balco, G., Schaefer, J., 2006. Cosmogenic-nuclide and varve chronologies for the deglaciation of southern New England. Quaternary Geochronology 1, 15–28.

Balco, G., Stone, J., Porter, S., Caffee, M., 2002. Cosmogenic-nuclide ages for New England coastal moraines, Martha's Vineyard and Cape Cod, Massachusetts, USA. Quaternary Science Reviews 21, 2127–2135.

Barrows, T., Stone, J., Fifield, L., Cresswell, R., 2002. The timing of the Last Glacial Maximum in Australia. Quaternary Science Reviews 21, 159–173.

Benson, L., Madole, R., Kubik, P., McDonald, R., 2007. Surface-exposure ages of Front Range moraines that may have formed during the Younger Dryas, 8.2 ka, and Little Ice Age events. Quaternary Science Reviews 26, 1638–1649.

Benson, L., Madole, R., Phillips, W., Landis, G., Thomas, T., Kubik, P., 2004. The probable importance of snow and sediment shielding on cosmogenic ages of north-central Colorado Pinedale and pre-Pinedale moraines. Quaternary Science Reviews 193, 193–206.

Bevington, P., Robinson, D., 1992. Data Reduction and Error Analysis for the Physical Sciences. WCB McGraw-Hill.

Cerling, T., Craig, H., 1994. Geomorphology and in-situ cosmogenic isotopes. Annual Reviews of Earth and Planetary Sciences 22, 273–317.

Clark, P., Brook, E., G.M., R., Yiou, F., Clark, J., 2003. Cosmogenic ^{10}Be ages of the Saglek Moraines, Torngat Mountains, Labrador. Geology 31, 617–620.

Gosse, J. C., Evenson, E., Klein, J., Lawn, B., Middleton, R., 1995. Precise cosmogenic ^{10}Be measurements in western North America: support for a global Younger Dryas cooling event. Geology 23, 877–880.

Gosse, J. C., Phillips, F. M., 2001. Terrestrial in situ cosmogenic nuclides: theory and application. Quaternary Science Reviews 20, 1475–1560.

Hallet, B., Putkonen, J., 1994. Surface dating of dynamic landforms: young boulders on aging moraines. Science 265 (5174), 937.

Phillips, F., Zreda, M., Benson, L., Plummer, M., Elmore, D., Sharma, P., 1996. Chronology for fluctuations in late Pleistocene Sierra Nevada glaciers and lakes. Science 274, 749–751.

Putkonen, J., Swanson, T., 2003. Accuracy of cosmogenic ages for moraines. Quaternary Research 59, 255–261.

Rinterknecht, V., Clark, P., G.M., R., Yiou, F., Bitinas, A., Brook, E., Marks, L., Zel$\breve{c}$s, V., Lunkka, J.-P., Pavlovskaya, I., Piotrowski, J., Raukas, A., 2006. The last deglaciation of the southeastern sector of the scandinavian ice sheet. Science 311, 1449–1452.

Rinterknecht, V., Clark, P., Raisbeck, G., Yiou, F., Brook, E., Tschudi, S., Lunkka, J.-P., 2004. Cosmogenic ^{10}Be dating of the Salpausselk$\ddot{a}$ I moraine in southwestern Finland. Quaternary Science Reviews 23, 2283–2289.

Schaefer, J., Denton, G., Barrell, D., Ivy-Ochs, S., Kubik, P., Andersen, B., Phillips, F., Lowell, T., Schlüchter, C., 2006. Near-synchronous interhemispheric termination of the Last Glacial Maximum at mid-latitudes. Science 312, 1510–1513.

Schaefer, J., Tschudi, S., Zhao, Z., Wu, X., Ivy-Ochs, S., Wieler, R., Baur, H., Kubik, P., Schluüchter, C., 2002. The limited influence of glaciations in Tibet on global climate over the past 170,000 yr. Earth and Planetary Science Letters 194, 287–297.

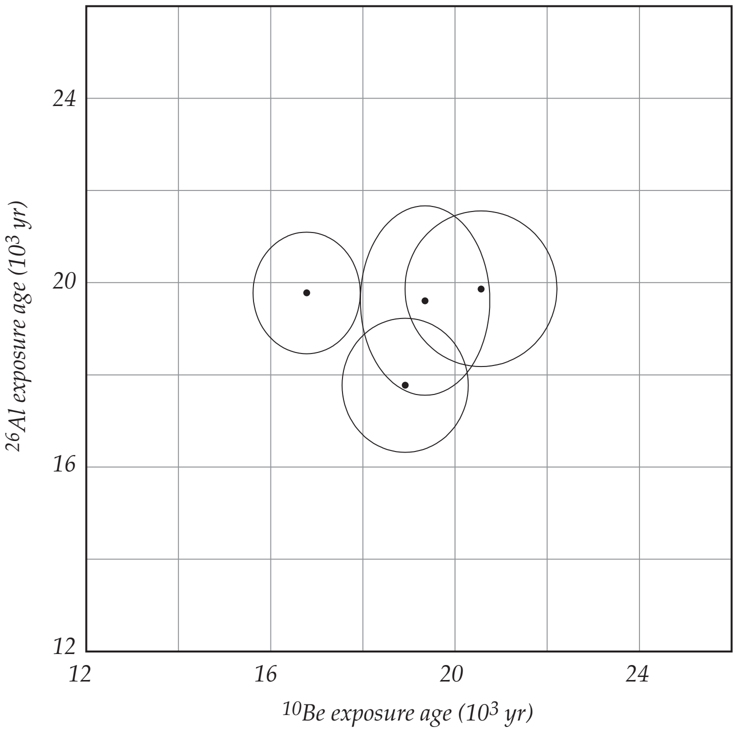

Fig. 1. ^{10}Be and ^{26}Al exposure ages from the Litang moraine, Tibet (Schaefer et al., 2002). Here and in succeeding figures, i) the ellipses show 68% probability regions and reflect measurement uncertainties only, and ii) these exposure ages differ from those in the source paper – for consistency, all ages have been recalculated from the source data using version 1 of the CRONUS-Earth exposure age calculator (http://hess.ess.washington.edu/math/). The ^{26}Al measurements in this particular study (Schaefer et al., 2002) are normalized to a different isotope ratio standard than the production rates in the CRONUS-Earth online calculators, so the ^{26}Al exposure ages shown in this figure may be inaccurate. This does not affect the correlation between ^{10}Be and ^{26}Al ages, so it is not important for the present purpose. Likewise, in this and succeeding examples, all the samples in the data set are close together, so the choice of local surface production rates or production rate scaling methods does not affect the scatter in the measurements or affect the correlation between ^{26}Al and ^{10}Be exposure ages.

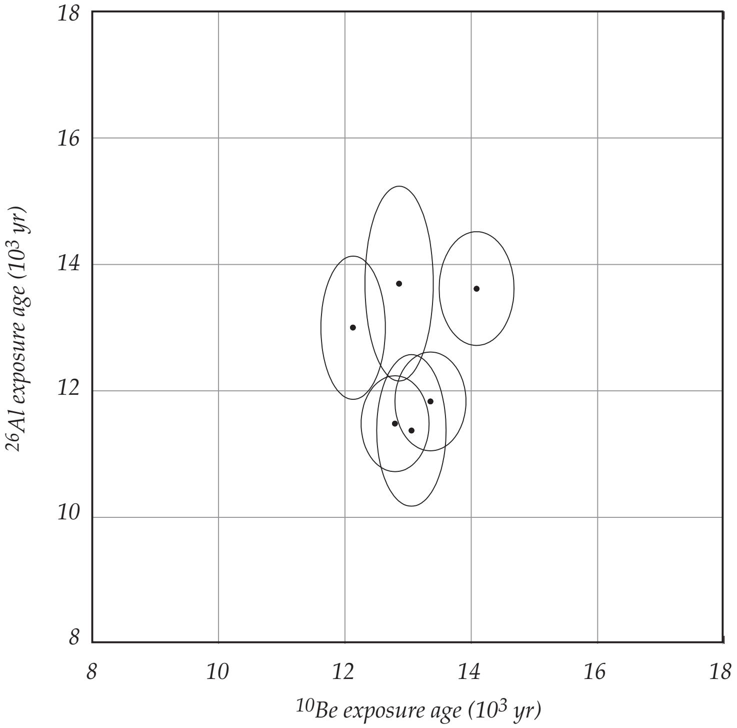

Fig. 2. ^{10}Be and ^{26}Al exposure ages from the Titcomb Lakes moraine (Gosse et al., 1995) .

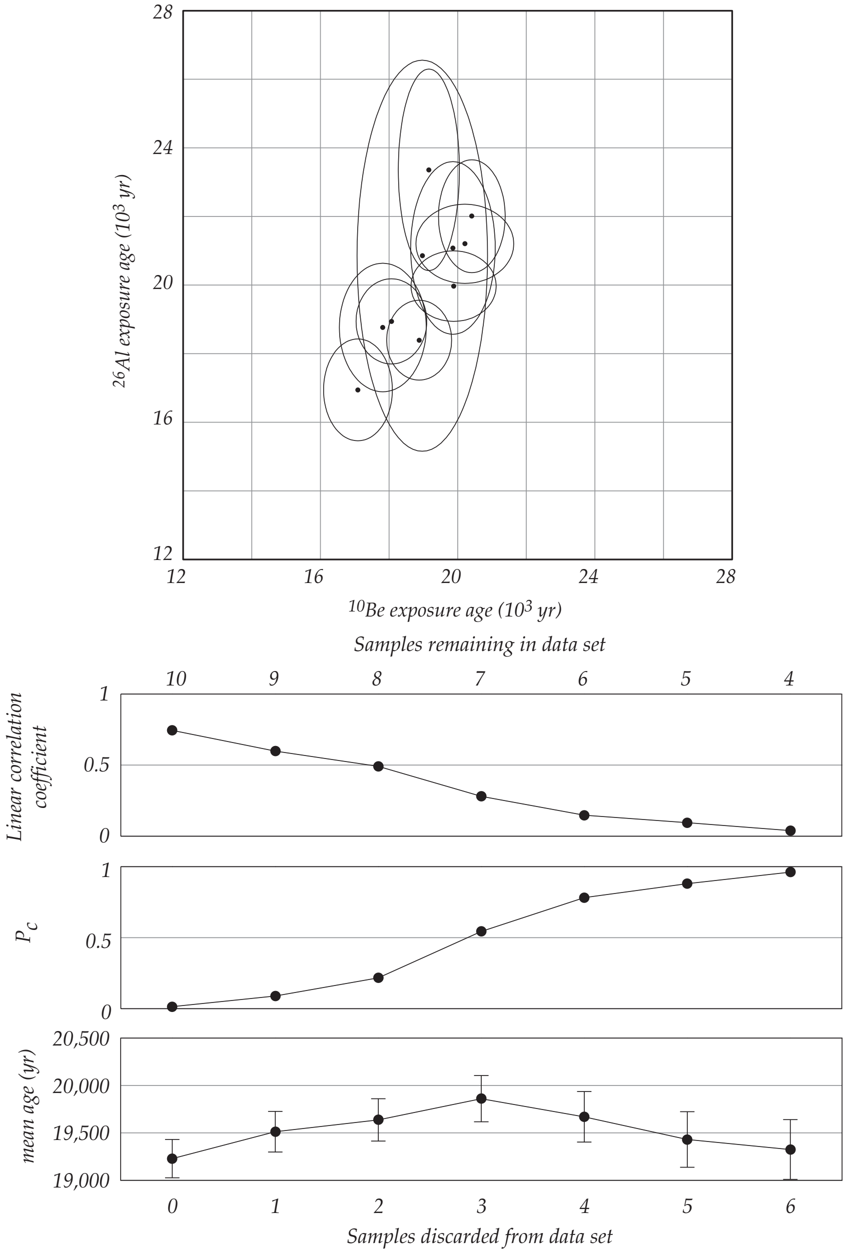

Fig. 3. Upper panel, ^{10}Be and ^{26}Al exposure ages from the Buzzards Bay moraine (Balco et al., 2002). Lower panel, results of pruning the data set as described in the text. The plot denoted 'mean age' shows the error-weighted mean and 1-σ uncertainty of the remaining measurements.

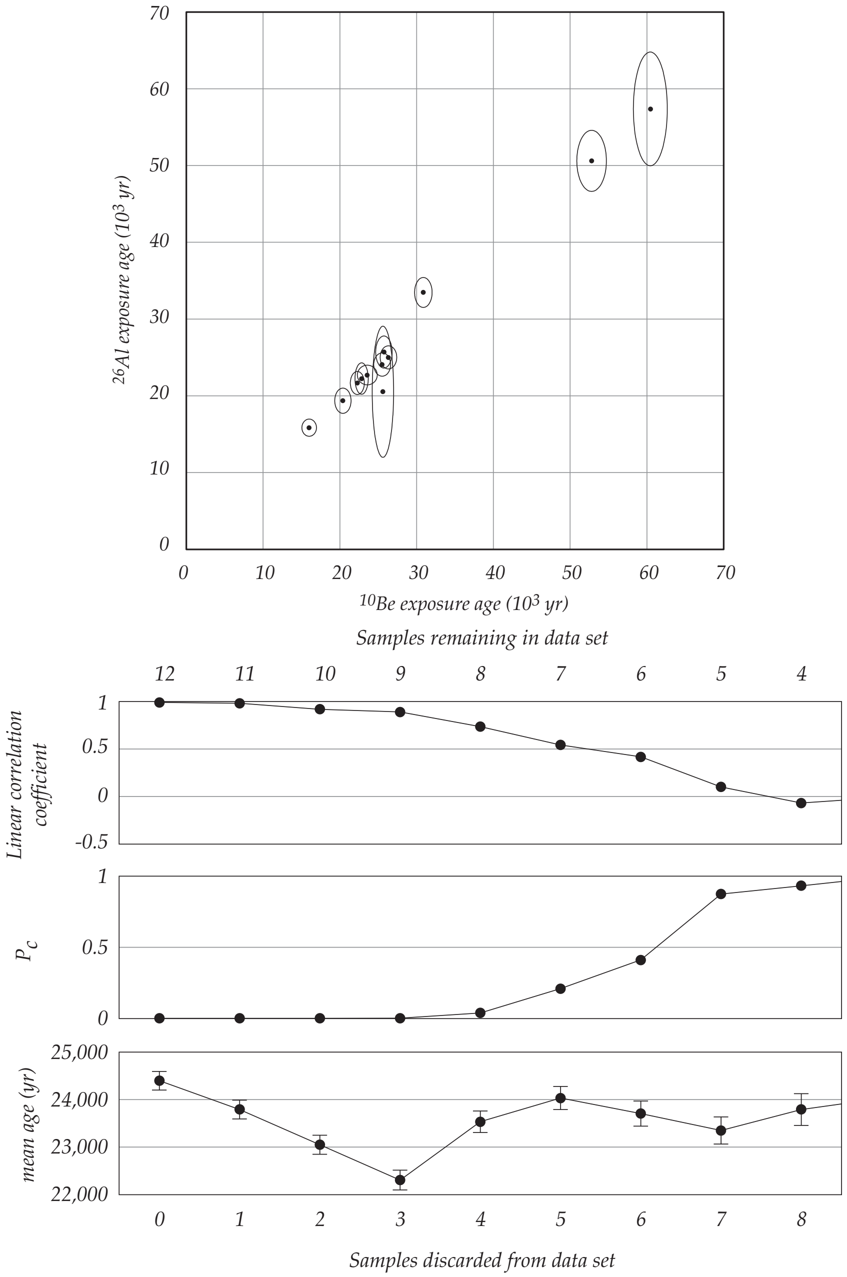

Fig. 4. Upper panel, ^{10}Be and ^{26}Al exposure ages from the Martha's Vineyard moraine (Balco et al., 2002). Lower panel, results of pruning the data set as described in the text. The plot denoted 'mean age' shows the error-weighted mean and 1-σ uncertainty of the remaining measurements.