NEW ZEALAND JOURNAL OF ARCHAEOLOGY

Computer Programs for Creating Cumulative Probability Curves and Annual Frequency Distribution Diagrams with Radiocarbon Dates

Tom Dye[1] and Eric Komori[2]

ABSTRACT

Statistical graphics for displaying aggregate ^{14}C data include the histogram, bar chart, and cumulative probability curve. An alternative graphic, the annual frequency distribution diagram, is described. The graphical limitations of the histogram and bar chart are well known and a potential problem with the temporal interpretation of cumulative probability curves exists. The annual frequency distribution diagram is free of these limitations and potential problem. The construction of annual frequency distribution diagrams and bootstrap confidence intervals is described. Computer programs for constructing cumulative probability curves and annual frequency distribution diagrams with bootstrap confidence intervals are made available for general use.

Keywords: RADIOCARBON, STATISTICAL GRAPHICS, COMPUTER PROGRAMS, ANNUAL FREQUENCY DISTRIBUTION DIAGRAM.

INTRODUCTION

Radiocarbon dates are an indispensable tool for archaeologists working with the space/time systematics of prehistory. They are used primarily to give absolute dates to events that otherwise could be placed only in relative order, a fact that underscores the importance of the dated event in prehistoric interpretation. Recently, however, as the number of radiocarbon dates available for analysis has grown, radiocarbon dates have begun to be analysed on their own as aggregate data sets whose chronological distribution can be used as a basis for inferences on the duration and intensity of some cultural trait. In the Pacific, three statistical graphics have been used to portray the chronological distributions of aggregate radiocarbon dates: the histogram, its close relative the bar chart, and the cumulative probability curve. After pointing out some limitations of the histogram and the bar chart, and a potential difficulty with the temporal interpretation of the cumulative probability curve, we describe an alternative statistical graphic, which we call the annual frequency distribution diagram. The annual frequency distribution diagram is a logical extension of the histogram, bar chart, and cumulative probability curve, whose properties make it preferable for certain types of analyses. A set of computer programs that construct cumulative probability curves and annual frequency distribution diagrams from the output

[1]State of Hawaii, Department of Land and Natural Resources, Historic Preservation Division, 33 South King street, 6th Floor, Honolulu, Hawaii 96813
[2]Bernice P. Bishop Museum, P.O. Box 19000A, Honolulu, Hawaii 96817-0916

New Zealand Journal of Archaeology, 1992, Vol. 14, pp. 35–43.

of the CALIB program (Stuiver and Reimer 1986) are described and made available for general use.

GRAPHICAL REPRESENTATIONS OF AGGREGATE ^{14}C DATA

The simplest graphical representations of aggregate ^{14}C data are the histogram and bar chart, both of which are widely used to summarise data distributions (Chambers *et al.* 1983: 24 ff.). The bar chart has a history that stretches back to the development of statistical graphics in the eighteenth century (Tufte 1983: 33). Histograms and bar charts are constructed by dividing the range of data into equal intervals, counting the number of observations in each interval, and plotting the counts as bar lengths. Histograms with interval widths of 200 years and bar charts with interval widths of 100 years have recently been used to portray the chronological distribution of calibrated radiocarbon dates from archaeological sites in the Pacific (Hunt and Holsen 1991; Kirch and Hunt 1988). This approach has a number of drawbacks. First, both displays reduce the information in the data by combining disparate observations in the same interval. In practice, the choice of a suitable interval width is governed by the amount of accuracy that can be tolerably lost; the wider the interval the greater the loss of accuracy (Cleveland 1985: 125). Second, the placement of the intervals, routinely determined by the desire to have intervals range between two round numbers, can affect the appearance of a histogram or a bar chart and lead to quite different impressions about the distribution of the data (Chambers *et al.* 1983: 24 ff.). Third, the bar corners are determined by the choice of interval width, but convey little information about the data. Finally, and perhaps most important, both the simple histogram and the bar chart ignore information carried in the statistical uncertainty (the ± term) of the radiocarbon age. A wide variety of statistical graphics has been developed to overcome these and other graphical limitations of the histogram and the bar chart (Chambers *et al.* 1983), and the problem of ignoring the statistical uncertainty of the radiocarbon age was partially overcome by Geyh, who used a histogram approximation of a Gaussian curve to add in the effects of statistical uncertainty for each sample (Geyh 1980; Geyh and Maret 1982).

A second graphical representation of aggregate ^{14}C data was developed by Law for Black and Green (1977). Law described the displays he produced as smooth frequency curves, but they were later named cumulative probability curves by Anderson (1989), whose displays were constructed by Foss Leach. The first step in their construction uses the radiocarbon age and statistical uncertainty to construct a Gaussian curve for each age determination. The Gaussian curve is used because it approximates the expected radiocarbon age distribution of multiple determinations for a single radiocarbon sample (Taylor 1987: 103, Fig. 4.15) and thus takes into account the statistical nature of the radiocarbon dating process. Gaussian curves for a suite of age determinations are summed to produce an aggregate curve whose smoothness depends on the interval width. The apparent smoothness of Anderson's aggregate curves is due to the use of an effective interval width of one radiocarbon year and an x-axis scale several hundreds of radiocarbon years wide. Law's curves, which span a period of over 3 millennia, appear smooth with an interval width of ten years. Cumulative probability curves, which incorporate the information in the statistical uncertainty of the radiocarbon age, also overcome the graphical limitations of the simple histogram and the bar chart.

Constructed in this way, a cumulative probability curve shows the distribution of a set of radiocarbon dates over the radiocarbon time scale. The shape of this curve is often interesting. Law pointed out that under the proper conditions "the distribution may suggest the intensity of prehistoric occupation through time" (Black and Green 1977) and Anderson used observations on skewing to infer that moa hunting intensified more slowly than it died out. A difficulty with these and other inferences based on cumulative probability curves is that radiocarbon years and calendar years are not equal, nor can their relationship be described with a constant. A large body of work dating materials of known age has shown that the relationship between radiocarbon years and calendar years varies over time, primarily because of fluctuations in the production of ^{14}C in the upper atmosphere. Thus, apart from very general statements, the temporal interpretation of inferences based on cumulative probability curves is unclear. Using Anderson's inference as an example, it may have taken a greater number of radiocarbon years for moa hunting to intensify than to die out, but the variable relationship between radiocarbon and calendar years makes it impossible to infer, on the basis of cumulative probability curves alone, that intensification took a greater number of *calendar* years.

This potential difficulty with the temporal interpretation of cumulative probability curves led us to investigate an alternative statistical graphic based on output from CALIB, a computer program that converts radiocarbon ages to calendar years. Recent versions of the CALIB program provide an option that calculates relative probability curves for individual age determinations. This option uses information on the relationship of radiocarbon ages and calendar years to modify the shape of the Gaussian curve for each sample. The amount of modification to the Gaussian curve is, in part, a function of the age of the dated material. In extreme cases, which are particularly common for materials dating to the last 300 years, two or more calendar years share the same radiocarbon age, so that the first-order relative probability curve produced by CALIB is multi-modal. Older materials generally yield a unimodal first-order curve, but with numerous second-order peaks and valleys. Figure 1 compares three graphical representations of a single radiocarbon age determination; first as its contribution to a histogram, second as a Gaussian curve centered at the radiocarbon age B.P., and last as output from CALIB. As the CALIB output for this example shows, a Gaussian curve may be a poor approximation of the calendrical age probability distribution, especially over the last 300 years.

THE ANNUAL FREQUENCY DISTRIBUTION DIAGRAM

The mechanics of constructing an annual frequency distribution diagram are nearly identical to those of the cumulative probability curve. The two displays differ primarily in the nature of the data used as input. CALIB relative probability curves, the raw data from which an annual frequency distribution diagram is constructed, are calculated by the CALIB program so that the maximum relative probability for any year is unity. This means that the areas beneath curves can differ from one another. In order to ensure that each sample contributes equally to the annual frequency distribution diagram, the relative probability curve for each sample is first normalised so that the sum of probabilities in its range is unity. These normalised curves are then summed.

An estimate of the systematic error in this procedure can be gained by constructing an annual frequency distribution diagram for a known distribution and comparing observed and expected values. Figure 2 shows two annual frequency distribution diagrams constructed from simulated samples. The sample data for each annual frequency distribution diagram consisted of one radiocarbon event annually, to produce a uniform distribution over a range

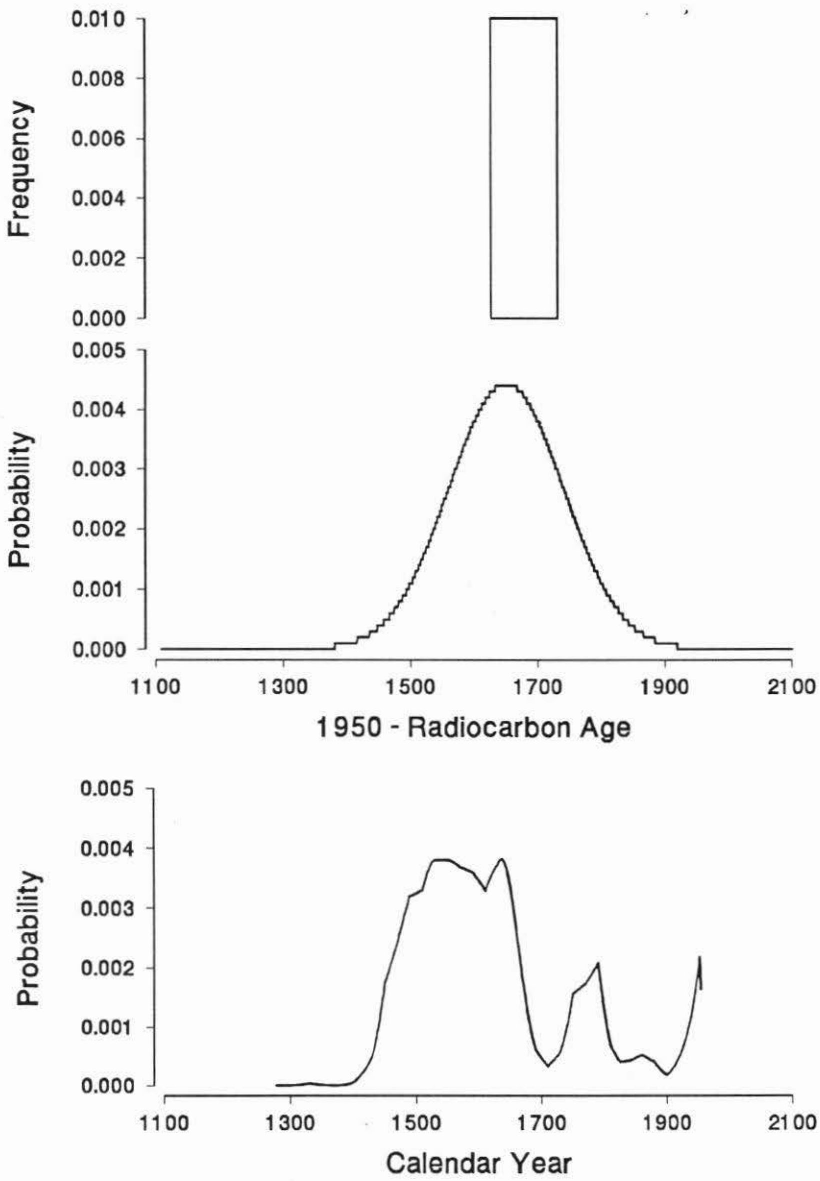

Figure 1: Three graphical representations of a single datum: a) histogram; b) Gaussian curve; and c) CALIB relative probability curve normalised so the area under the curve is unity. The jagged portions of the Gaussian curve are due to the file format of CALIB output files, which records numbers to four decimal places. Extra decimal places would yield a smoother curve. Note the approximate Gaussian shape of the left-hand side of the CALIB relative probability curve and the extreme deviations of the right-hand side. The deviations are due to the pronounced deVries effects for the period after A.D. 1650, and to the Suess effect after A.D. 1900 (see Taylor 1987: 35 ff.).

encompassing the interval from 6000 B.C. to A.D. 1950. The over 8,000 events in each sample were converted to conventional radiocarbon ages. The radiocarbon ages in the first sample were assigned an experimental uncertainty of 50 radiocarbon years and those in the second sample an experimental uncertainty of 100 radiocarbon years. The radiocarbon ages were then treated as described above to produce Figure 2. The figure shows that the observed distribution of each sample is centered around the expected distribution, with a mean of 1.0, and that deviations from the expected distribution vary inversely with the size of the mean experimental uncertainty. The observed distribution of the sample with an experimental uncertainty of 50 ranges between 0.46 and 1.24, with a standard deviation of 0.1. The observed distribution of the sample with an experimental uncertainty of 100 ranges between 0.75 and 1.12, with a standard deviation of 0.06. In our experience, the mean experimental uncertainty in large radiocarbon datasets is close to 100, so that the small standard deviation of this latter sample lends credence to the claim that the procedures used to construct an annual frequency distribution diagram yield curves that are analytically useful.

Systematic error, even of this relatively small magnitude, has the potential to confound analysis, so it is useful to normalise an annual frequency distribution diagram for this source of error. In general, this can be accomplished by calculating the percent deviation from the mean at each year of a curve like Figure 2 and subtracting this percentage of the y-axis

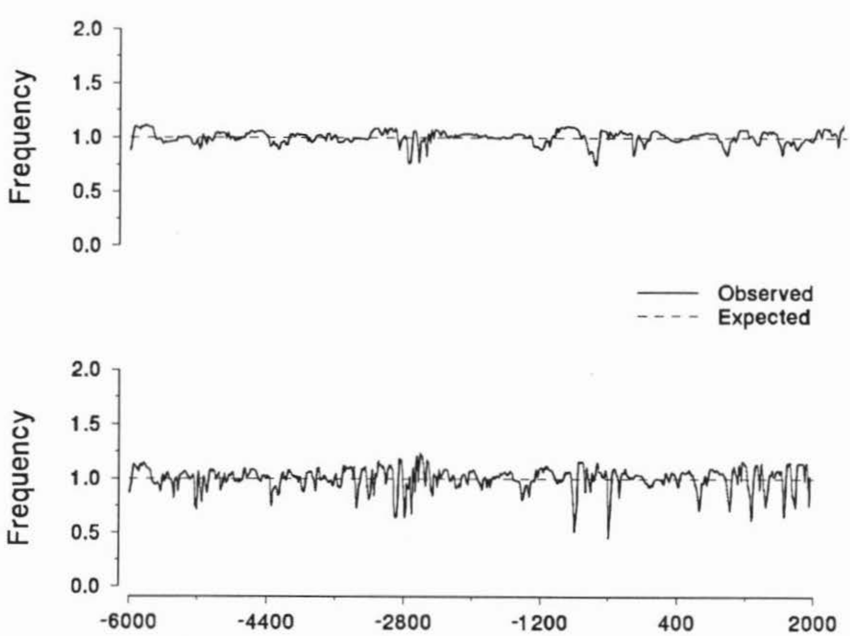

Figure 2: Observed and expected values for two annual frequency distribution diagrams of hypothetical samples of known distribution. The upper sample has an experimental uncertainty of 100 ^{14}C years; the lower sample 50 ^{14}C years. Note the relative smoothness of the upper curve.

height from the corresponding year of an annual frequency distribution diagram. Ideally, distributions with the same general shape, statistical uncertainty, and size of each data set analysed would be generated, but this practice is limited by the capacity of computers generally available to prehistorians. Therefore, standard normalisation curves derived from uniform distributions can be used, with the choice of normalisation curve dependent upon the mean statistical uncertainty of the data being analysed. Clearly, this procedure will rarely remove all of the systematic error from an annual frequency distribution diagram and we discuss below constraints this might place on the interpretation of annual frequency distribution diagrams.

A more intuitive feel for the validity of the annual frequency distribution diagram construction procedure can be gained from Figure 3, which compares three graphical representations of the same data set. Here it can be seen that the histogram, the cumulative probability curve, and the annual frequency distribution diagram all share the same basic first order shape, with a long low tail to the left and a major peak followed by a decline at the right. Many of the details vary. Note particularly the differences between the slopes leading to and from the major peak of the cumulative probability curve and the annual frequency distribution diagram. Temporal inferences about the development and decline of the cultural trait associated with this major peak might be affected by the choice of graphical display.

CONFIDENCE INTERVALS

In most situations, annual frequency distribution diagrams will be constructed from a sample drawn from a population of potential radiocarbon age determinations. When this is the case, and the distribution of the population, rather than the sample, is of interest, then it is desirable to compute confidence intervals for the annual frequency distribution diagram. Recent advances in statistical method, made possible through the falling costs and greater availability of computer time, now provide a means for calculating confidence intervals for data, like those of aggregate ^{14}C age determinations, that do not conform to known, or mathematically analysed, distributions. We have chosen one of these, the bootstrap, to calculate confidence intervals directly from the data, without assumptions about their distribution (Efron 1979; Efron and Tibshirani 1991).

The bootstrap is built on the concept of the bootstrap sample, which is a sample of size n drawn with replacement from a data set $x = (x_1, x_2, ..., x_n)$. In a bootstrap sample $x^* = (x^*_1, x^*_2, ..., x^*_n)$ each x^*_n is a randomly selected member of the original data set. In rare instances $x = x^*_n$, but it is more usually the case that some x_i are not selected and others are selected more than once. Intervals with a desired confidence C are computed by drawing a large number B of independent bootstrap samples, each of size n, with a random number generator; constructing an annual frequency distribution diagram for each of the B bootstrap samples; and finding, for each year of the annual frequency distribution diagram range, the empirical interval within which C samples fall. Where $C = 0.9$, the minimum number of bootstrap samples required is 20; the smallest and largest values at each year are discarded and the remaining range of 18 samples ($^{18}/_{20} = 0.9$) yields the bootstrap confidence interval.

Figure 4 shows the data set of Figure 3 with .95 bootstrap confidence intervals. The confidence intervals are sufficiently wide to encompass any systematic error that was not removed by the normalisation procedure. In fact, the width of the intervals for a sample this size obscures all but the first order shape of the curve, with its peak between A.D.

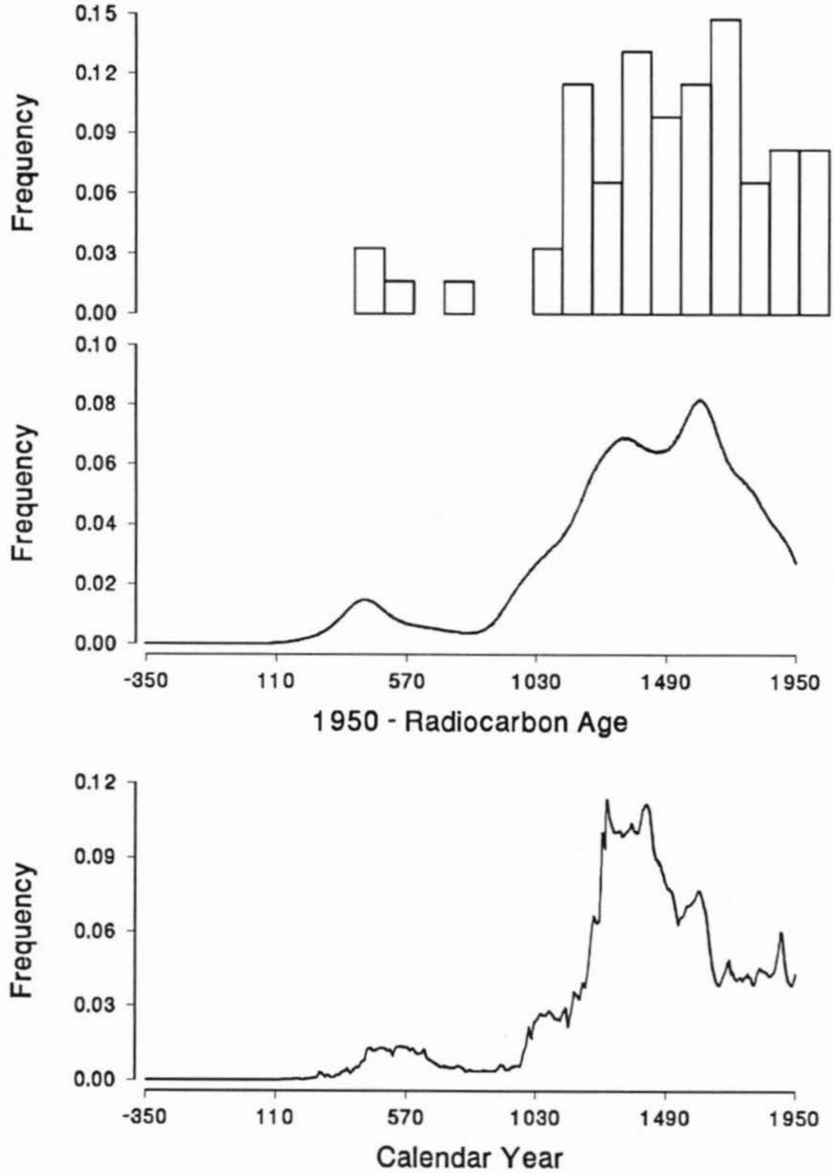

Figure 3: Three graphical representations of a single data set: a) frequency histogram; b) cumulative probability curve; and c) annual frequency distribution diagram. The data are 61 ^{14}C dates from O'ahu, Hawai'i; they are used for illustrative purposes only. All three graphical representations share a similar first order shape: a long, low tail at the early end of the x-axis is followed by a peak toward the late end and a subsequent decline. The details vary somewhat. Note the different shapes of the major peak: the histogram and the cumulative probability curve both yield relatively wide peaks, with fairly gentle left-hand slopes. In contrast, the slope on the left-hand side of the annual frequency distribution diagram is abrupt.

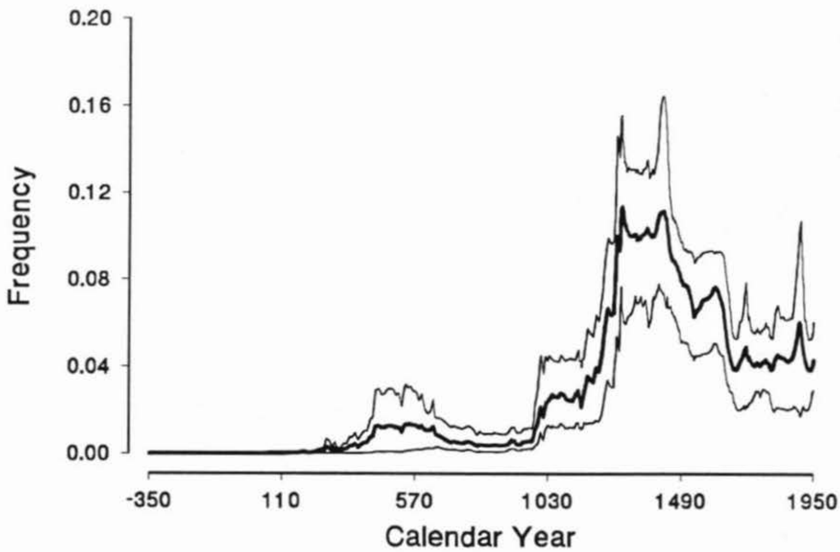

Figure 4: The data set of Figure 3, with 0.95 bootstrap confidence intervals.

1300–1450 and subsequent decline. In our own work with annual frequency distribution diagrams we resist the temptation to invest meaning in small-scale features of the curve. Instead, we visualise the smoothest curve that can be drawn within the constraints of the confidence intervals, and limit our inferences to these relatively large-scale features of the curve. Interpreted this way, we are confident that the small amount of residual systematic error in an annual frequency distribution diagram does not confound analyses.

TWO COMPUTER PROGRAMS: *SUM^{14}C* AND *UNCALIB*

In the course of our work with aggregate ^{14}C data we developed two computer programs, SUM^{14}C and UNCALIB, to facilitate the construction of cumulative probability curves, annual frequency distribution diagrams, and bootstrap confidence intervals. UNCALIB uses information on the radiocarbon age and statistical uncertainty of dated materials to produce the Gaussian curves used in the construction of cumulative probability curves. SUM^{14}C sums Gaussian or CALIB relative probability curves to produce either cumulative probability curves or annual frequency distribution diagrams, and will construct bootstrap confidence intervals for both. UNCALIB and SUM^{14}C, for use with the DOS 2.0 and greater, are both available either from the senior author or from the anonymous FTP site 155.68.2.2 in the SAS sub-directory (Society for Archaeological Sciences).

REFERENCES

Anderson, A. 1989. Mechanics of overkill in the extinction of New Zealand moas. *Journal of Archaeological Science* 16: 137–151.

Black, S. J. and Green, R. C. 1977. Radiocarbon dates from the Solomon Islands to 1975. *Oceanic Prehistory Records* 4: 1–56. With a contribution by Garry Law.

Chambers, J. M., Cleveland, W. S., Kleiner, B. and Tukey, P. A. 1983. *Graphical methods for data analysis*. Wadsworth and Brooks/Cole Statistics/Probability Series. Wadsworth and Brooks/Cole.

Cleveland, W. S. 1985. *The elements of graphing data*. Wadsworth Advanced Books and Software.

Efron, B. 1979. Bootstrap methods: Another look at the jackknife. *Annals of Statistics* 7: 1–26.

Efron, B. and Tibshirani, R. 1991. Statistical data analysis in the computer age. *Science* 253: 390–395.

Geyh, M. A. 1980. Holocene sea-level history: a case study in the statistical evaluation of ^{14}C dates. *Radiocarbon* 22: 695–704.

Geyh, M. A. and de Maret, P. 1982. Histogram evaluation of ^{14}C dates applied to the first complete Iron Age sequence from west central Africa. *Archaeometry* 24: 158–163.

Hunt, T. L. and Holsen, R. M. 1991. An early radiocarbon chronology for the Hawaiian Islands: a preliminary analysis. *Asian Perspectives* 30: 147–161.

Kirch, P. V. and Hunt, T. L. (Eds) 1988. *Archaeology of the Lapita cultural complex: a critical review*. Thomas Burke Washington State Museum Research Report 5. Thomas Burke Memorial Washington State Museum.

Stuiver, M. and Reimer, P. J. 1986. A computer program for radiocarbon age calibration. Proceedings of the 12th International ^{14}C Conference, *Radiocarbon* 28 (2B): 1022–1030.

Taylor, R. E. 1987. *Radiocarbon dating: an archaeological perspective*. Academic Press.

Tufte, E. R. 1983. *The visual display of quantitative information*. Graphics Press.

Received 10 June 1992
Accepted 15 December 1992