[RADIOCARBON, VOL. 35, NO. 2, 1993, P. 317-322]

A SIMPLIFIED APPROACH TO CALIBRATING ^{14}C DATES

A. S. TALMA and J. C. VOGEL

Quaternary Dating Research Unit, CSIR, P. O. Box 395, Pretoria 0001, South Africa

ABSTRACT. We propose a simplified approach to the calibration of radiocarbon dates. We use splines through the tree-ring data as calibration curves, thereby eliminating a large part of the statistical scatter of the actual data points. To express the age range, we transform the ± 1 σ and ± 2 σ values of the BP age to calendar dates and interpret them as the 68% and 95% confidence intervals. This approach bypasses the conceptual problems of the transfer of individual probability values from the radiocarbon to the calendar age. We have adapted software to make this calibration possible.

INTRODUCTION

The tree-ring calibration data now available have prompted the user community to insist increasingly on working with the calibrated ages of their ^{14}C dates. Different computer programs have been developed to accomplish conversion of ^{14}C (BP) ages to calendar dates (*e.g.*, Stuiver & Reimer 1986; van der Plicht & Mook 1989). Experience shows that considerable confusion can arise in the interpretation of values obtained from available programs (see review by Aitchison *et al.* 1989). For this reason, we propose a somewhat simplified calibration procedure without actually sacrificing accuracy. First, we suggest that the calibration curve be smoothed to a certain extent without seriously harming the real "wiggles" of the ^{14}C levels. Second, we propose that the counting error, expressed as ± 1 σ, should be transformed directly to a corresponding 68% confidence interval of calendar dates on the X-axis.

SMOOTHING OF THE CALIBRATION CURVE

^{14}C analyses of dated tree rings have provided data sets (x_i, y_i, σ_i), where

x_i= AD/BC date of the ring
y_i= ^{14}C BP age of the ring
σ_i= standard deviation of the BP age measurement.

The aim is to use such a data set to derive the most acceptable relation between y and x, $y = g(x)$, that can be used for interpolation of the measured BP ages (y) to obtain calibrated dates of samples (x). In effect, one seeks the best estimate of past atmospheric ^{14}C levels.

Although it can be argued that the most probable value of the ^{14}C level at a certain point in time (*sic*!) is represented by the actually measured ^{14}C age, it is also true that the real calibration curve should not pass through all of the measured points. In fact, counting statistics demand that, on average, 68.2% of the measured points should be within ± 1 σ of the actual curve, 27.2% of the data points in the 1–2 σ interval and the remaining 4.6% > 2 σ from the calibration curve. A curve produced by connecting all the measured points by straight lines or by a flowing curve will thus contain numerous small-scale fluctuations that realistically do not represent the actual variations of the past atmospheric ^{14}C level. The slow decay of the post-bomb atmospheric ^{14}C level indicates the sluggishness of the atmosphere/ocean system in this respect, and places an upper limit on the rate of short-term decrease of atmospheric ^{14}C levels.

Our work with a calibration data set of 1–3-year tree-ring samples (Vogel *et al.* 1993) has convinced us that smoothing techniques, such as multiple-point running averages, are unsatisfactory, and a more statistically justified approach is required.

In this situation, the most appropriate mathematical technique to apply is the spline curve (Reinsch 1967) with a "stiffness" that corresponds to the statistical uncertainty of the individual measured values. This approach also meets the geophysical requirement that the ^{14}C level in the atmosphere does not change too abruptly.

Splines are cubic interpolation curves, $g(x_i)$, specified separately for each interval between two adjacent measuring points: (x_i, y_i) and (x_{i+1}, y_{i+1}) (Reinsch 1967). They are calculated from the entire data set based on the conditions that

- The curve be continuous at the data points
- The integral slope change rate be a minimum, compatible with required precision constraints
- The variability, or smoothness, of the curve can be specified to match the precision of the analytical data.

The precision of the spline curve, s, is defined as

$$s^2 = \frac{\Sigma\,(g(x_i)-y_i)/\sigma_i)^2}{n} \quad .$$

This is a parameter of the closeness of fit or, inversely, the stiffness of the curve. Calculating $g(x_i)$ with $s = 0$ implies that the curve thus produced connects each of the data points. This cannot be a valid representation of the real-world situation, because it suggests that each data point be true, an unrealistic requirement in view of the counting statistics. Setting s to a large number allows maximum deviation between the line and the data points and will produce a near-straight line. Stipulating $s = 1$ is the most realistic requirement, because it produces a curve, such that the root-mean-square distance of the sample points is 1 σ away from the curve. Thus, it is the best compromise between the slowest varying calibration function and the counting statistics.

To illustrate, the example in Figure 1 shows a particularly rough part of the 2500 BC–AD 1940 calibration curve (Stuiver & Pearson 1986; Pearson & Stuiver 1986). The spline curve eliminates the "zig-zags" of the calibration curve while retaining the essential structure of the wiggles (*e.g.*, at 1400–1300 BC). Dates calibrated near a "zig-zag" in the curve produced by connecting all the points will have three or more possible calendar dates, whereas the spline curve will produce only one. This particular spline curve closely fits the requirements laid down by the counting statistics, *viz.*, 66.8% of the tree-ring measurements are within ± 1 σ of the spline curve, 28.3% in the 1–2 σ interval, and 4.9% beyond 2 σ. These percentages indicate that the spline is a good approximation of the past ^{14}C level and that deviations of the data points from the spline closely obey Gaussian statistics. Other smoothing techniques, such as running averages, rely on arbitrary choices of width or weighting that are selected on the basis of the final result for a given data set. Only splines can produce smoothing based on statistical requirements.

A useful and practical property of spline-curve calculation is that the x values of the data points need not be evenly spaced (Reinsch 1967), a considerable advantage when using irregularly spaced short-interval data (*e.g.*, Vogel *et al.* 1993).

Transformation of Probabilities

The purely Gaussian distribution of probability around the measurement of a ^{14}C sample (BP) is quantified by the 1 σ error attached to the age. The transformation of this probability distribution to the calendar date (or x-axis) is a complicated matter, and the result depends, to some extent, on the statistical approach taken (Dehling & van der Plicht 1993).

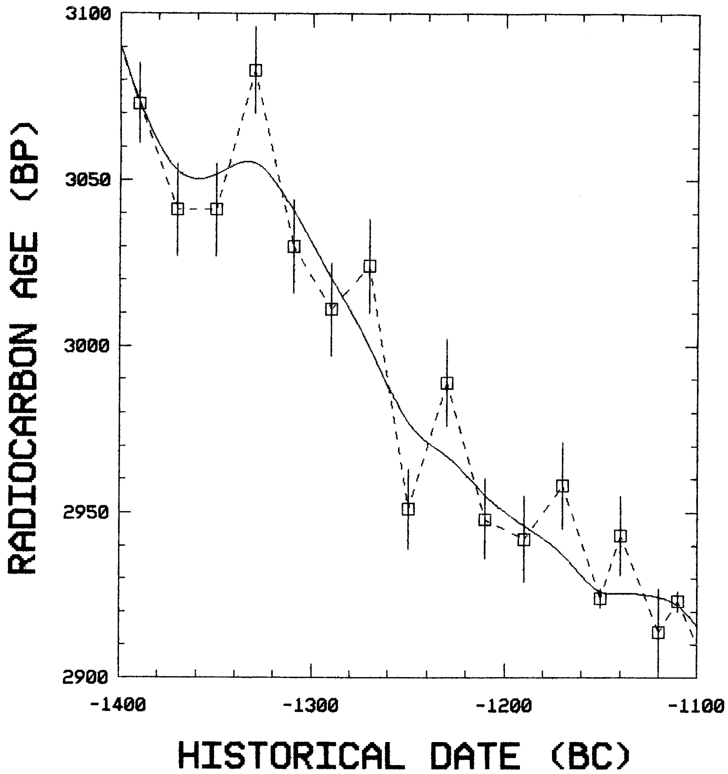

Fig.1. A selection of tree-ring calibration data from Pearson & Stuiver (1986) with the calculated spline curve. Note the manner in which the wiggle valley at 1350 BC is preserved by the spline curve. In the range, 3000–2900 BP, the spline curve averages between varying measurement points and removes the multiple intercepts that would have occurred if all the sample points were connected.

The textbook approach proceeds from a sample analysis as a measurement with a Gaussian probability distribution around the BP age. Projection of the probability density via the calibration curve onto the calendar date (X) axis is then required. This should produce a probability distribution around the calendar date. The alternative approach assumes that some distribution of calendar dates (prior knowledge) exists, and that the counting statistics reflect the (transformed) probability of such a distribution of occurrence. Its practical application is that probabilities are directly transformed from the Y- to X-axis. Differences between these approaches are evident where the calibration curve exhibits sharp changes in slope (Stuiver & Reimer 1989; Dehling & van der Plicht 1993). In our view, prior knowledge of an age distribution cannot be assumed when calibrating results of a ^{14}C analysis.

Whatever the merits of the various approaches, users of ^{14}C dates are accustomed to the concept of a ^{14}C age and of age ranges with 68% (or 95%) confidence levels. These can be converted legitimately into calibrated dates and their associated 68% (or 95%) confidence ranges on the calendrical scale. This implies projecting the BP ages and their $\pm$ 1 (or $\pm$ 2 σ) ranges from the Y-axis onto the X-axis. The 5 BP values are thereby transformed to 5 (or more) calendar dates, and the same probability statements concerning the date ranges can be made as for the original BP ages.

PRETORIA CALIBRATION PROCEDURE

Test Sample:	Pta-xxxx	+ 2 σ	AD 1416
^{14}C age:	440 ± 40 BP	+ 1 σ	AD 1429
Calibration file:	cal2.cal	cal-age	AD 1443
	Stuiver & Pearson (1986),	− 1 σ	AD 1461
	northern hemisphere	− 2 σ	AD 1492

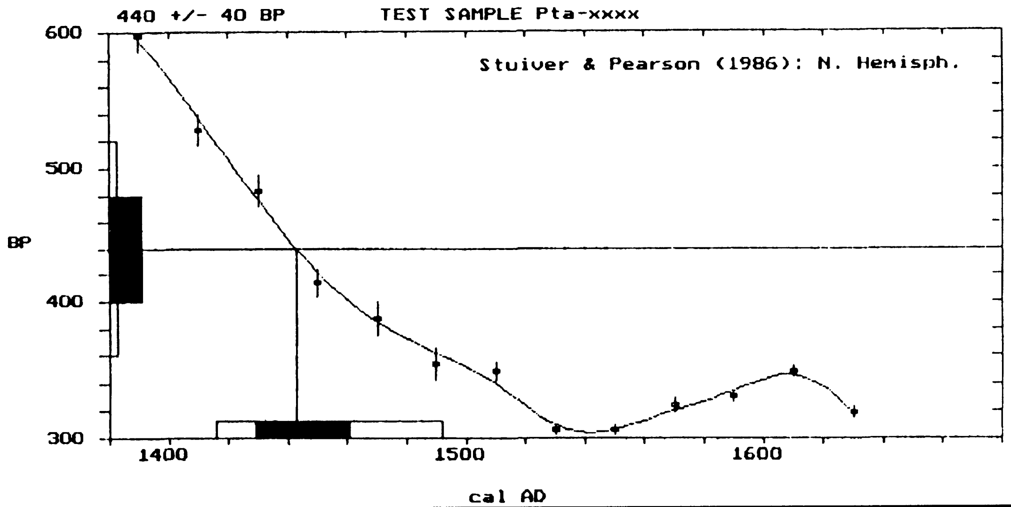

Fig. 2. Test example of calibration on a simple part of the calibration curve. Due to the change of slope, the calendar date ranges are distributed asymmetrically around the calibrated date.

This concept can be illustrated in a simple singular calibration case at a BP age of 440±40 (Fig. 2). At this point, only a single possible calendar date exists, but the 1 and 2 σ ranges are slightly asymmetrical due to the change in slope. In a more complex situation, where a single BP age corresponds to more than one calibrated date, the 1 and 2 σ ranges can be split, and more than one age range is possible (Fig. 3).

We have modified the Groningen calibration program, CAL4 (van der Plicht & Mook 1989), to accomplish the above calibration procedure (Figs. 2, 3). Different calibration data sets can be implemented to accommodate the reservoir effects of atmospheric ^{14}C in the northern and southern hemispheres (Stuiver & Pearson 1986; Pearson & Stuiver 1986; Vogel *et al.* 1993) and in the ocean (Stuiver, Pearson & Braziunas 1986). Our laboratory routinely provides these calibrated dates to sample submitters.

Sets and Series of ^{14}C Dates

Various ^{14}C laboratories have proposed procedures for dealing with groups of ^{14}C dates, and computer programs have been constructed for their evaluation (see review by Aitchison *et al.* 1989). The lack of a standard approach is bound to cause confusion in the user community; thus, further discussion is needed.

If a set of ^{14}C dates for a single event or a specific occupation level is available, the weighted mean of the ^{14}C results obviously can be transformed to a calendar date. The problem is that the

PRETORIA CALIBRATION PROCEDURE			
Test Sample:	Pta-xxxx	+ 2 σ	AD 1482
^{14}C age:	330 ± 20 BP	+ 1 σ	AD 1502
Calibration file:	cal2.cal	cal-age	AD 1516
	Stuiver & Pearson (1986),	− 1 σ	AD 1530
	northern hemisphere	− 2 σ	AD 1556
		cal-age	AD 1584
		cal-age	AD 1624
		− 1 σ	AD 1633
		− 2 σ	AD 1640

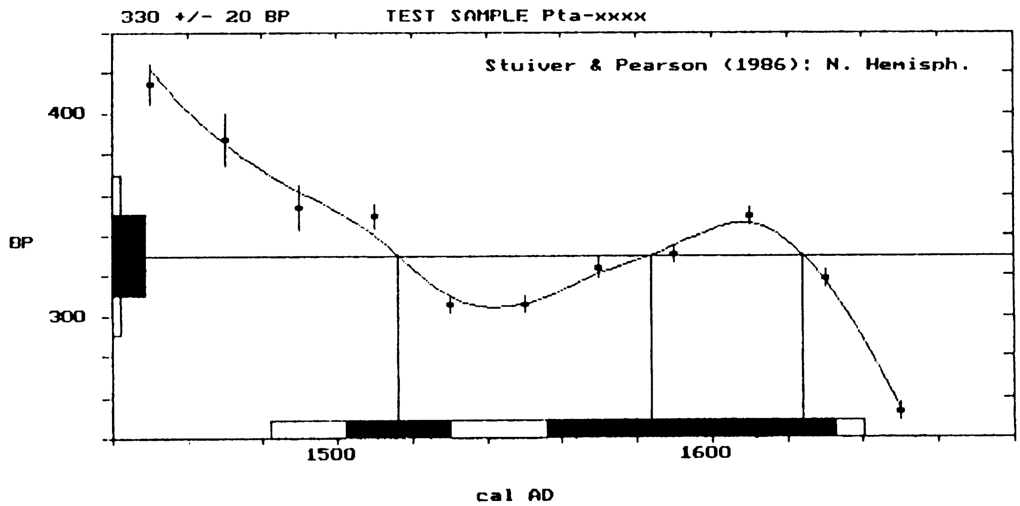

Fig. 3. Test example of a calibration at a multiple interception point on the calibration curve. Three calibrated ages are possible. Adding ± 1 σ to the BP age yields two separate age ranges, while ± 2 σ increases the age to a single wide range.

true contemporaneity of the samples is seldom certain. If it were, separate analyses would not be justified, and samples would be combined for a more accurate or longer measurement. In fact, the reason for producing more than one analysis on a specific level is that the sample need not be exactly the same age as the event with which it is associated. The most common problem (in the relevant time range) is that the sample, *e.g.*, charcoal, can contain older wood. For example, a set of four dates from a specific level should be enough to reveal such outliers, which can then be rejected. However, it is clear that such a set should not be averaged on any time scale. Waterbolk (1971) discussed this problem extensively and proposed procedures for interpreting both sets and series of dates. This approach should be applied to the individual dates after calibration.

Further, the "old wood" problem implies that ^{14}C dates are most useful for estimating the end of a depositional phase. The duration should be determined by dating the end of the preceding phase as well. A better estimate could be made in this way than, for example, in rejecting the outer quartiles, which would be rejecting half the measurements.

Thus, we believe that, in interpreting groups of ^{14}C dates, mathematical procedures are no longer very helpful, but insight and experience become more important.

Conclusions

We have attempted to show that the calibration data set should be considered with some degree of uncertainty, because it represents a set of measurements (with inherent analytical uncertainty) of past atmospheric ^{14}C levels. As such, some smoothing is justified, for which we have found the spline curve very useful. The spline calibration enables adjustment of the average curve by a quantified closeness-of-fit parameter to the measured data points, and reduces the number of situations where more than one calendar date is obtained by calibration.

In view of the complexity of transforming probabilities from the BP to the AD/BC axis, and the resulting confusion this creates among the ^{14}C user community, a simpler statement of 68% and 95% confidence intervals along the AD/BC axis seems to be the most practical current application of calibration programs.

Acknowledgments

We thank the following colleagues with whom we have communicated while formulating our concepts: J. van der Plicht in Groningen, C. Elphinstone in Pretoria and P. Reimer in Seattle. We thank M. Stuiver and J. van der Plicht for supplying us with their programs and calibration data sets.

References

Aitchison, T. C., Leese, M., Michczynska, D. J., Mook, W. G., Otlet, R. L., Ottoway, B.S., Pazdur, M. F., van der Plicht, J., Reimer, P. J., Robinson, S. W., Scott, E. M., Stuiver, M. and Weninger, B. 1989 A comparison of methods used for the calibration of radiocarbon dates. *In* Long, A. and Kra, R.S., eds., Proceedings of the 13th International ^{14}C Conference. ^{14}C 31(3): 846–864.

Dehling, H. and van der Plicht, J. 1993 Statistical problems in calibrating radiocarbon dates. *In* Stuiver, M., Long, A. and Kra, R. S., eds., Calibration 1993. *Radiocarbon* 35(1): 239–244.

Pearson, G. W. and Stuiver, M. 1986 High-precision calibration of the radiocarbon time scale, 500 BC–2500 BC. *In* Stuiver, M. and Kra, R. S., eds., Proceedings of the 12th International ^{14}C Conference. *Radiocarbon* 28(2B): 839–862.

Reinsch, C. H. 1967 Smoothing by spline functions. *Numerische Mathematik* 10: 177–183.

Stuiver, M. and Pearson, G. W. 1986 High-precision calibration of the ^{14}C time scale, AD 1950–500 BC. *In* Stuiver, M. and Kra, R. S., eds., Proceedings of the 12th International ^{14}C Conference. *Radiocarbon* 28(2B): 808–838.

Stuiver, M., Pearson, G. W. and Braziunas, T. F. 1986 Radiocarbon age calibration of marine samples back to 9000 cal yr BP. *In* Stuiver, M. and Kra, R. S., eds., Proceedings of the 12th International ^{14}C Conference. *Radiocarbon* 28(2B): 980–1021.

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

_____ 1989 Histograms obtained from computerized radiocarbon age calibration. *In* Long, A. and Kra, R. S., eds., Proceedings of 12th International ^{14}C Conference. *Radiocarbon* 31(3): 817–823.

van der Plicht, J. and Mook, W. G. 1989 Calibration of radiocarbon dates by computer. *In* Long, A. and Kra, R. S., eds., Proceedings of the 13th International ^{14}C Conference. *Radiocarbon* 31(3): 805–816.

Vogel, J. C., Fuls, A., Visser, E. and Becker, B. 1993 Pretoria calibration curve for short-lived samples, 1930–3350 BC. *In* Stuiver, M., Long, A. and Kra, R. S., eds., Calibration 1993. *Radiocarbon* 35(1): 73–85.

Waterbolk, H. T. 1971 Working with radiocarbon dates. *Prehistory Society Proceedings* 37: 15–33.