CALIBRATION OF RADIOCARBON AGES BY COMPUTER

JOHANNES VAN DER PLICHT and W G MOOK

Centre for Isotope Research, University of Groningen, Westersingel 34
9718 CM Groningen, The Netherlands

ABSTRACT. A PC-based computer program for automatic calibration of ^{14}C dates has been developed in Turbo-Pascal (version 4.0). It transforms the Gaussian ^{14}C dating result on the 3σ level into a real calendar age distribution. It uses as a calibration curve a spline function, generated along the calibration data points as published in the Radiocarbon Calibration Issue. Special versions of the code can average several ^{14}C dates into one calibrated result, generate smoothed curves by a moving average procedure and perform wiggle matching.

INTRODUCTION

In radiocarbon dating, variations in the atmospheric ^{14}C content complicate the conversion of conventional ages BP (*ie*, years before AD 1950) into real calendar ages (AD/BC) (de Vries, 1958; Willis, Tauber & Münnich, 1960). These variations are indirectly observed in tree rings from European and North American wood. During the last decades, hundreds of measurements have been made on dendrochronologically dated wood (Stuiver & Kra, 1986) resulting in the generally accepted calibration curves from Stuiver and Pearson (1986) (AD 1940 – 500 BC) and Pearson and Stuiver (1986) (500–2490 BC). The calibration points are given in intervals of 20, sometimes 15, years. These data are usually extended to 5210 BC, using the 10-yr interval data from Pearson *et al* (1986).

The calibration curve connecting these points has a highly irregular shape. This obviously complicates the calibration procedure, *ie*, the translation of a BP ^{14}C age into a real calendar age (cal AD/BC). For example, the ^{14}C age can correspond to more than one real calendar date due to medium term variations or wiggles (Suess, 1970).

The ^{14}C age results from a radioactive decay measurement; thus, it closely approaches statistically a Gaussian probability distribution with a well-defined standard deviation. The real calendar age probability distribution, however, is no longer Gaussian, unless the calibration curve happens to be a straight line in the time period considered (*ie*, there is a linear relationship between cal AD/BC and BP). The calibration age, then, cannot generally be stated as the most probable result, with an associated error bar.

The following procedure was adopted. First, the calibration curve uncertainty is taken into account according to

$$\sigma = \sqrt{(\sigma_{date}^2 + \sigma_{curve}^2)}$$

This is mathematically not strictly correct (it only applies to a straight line as a calibration curve), but the error made is not large. Next, an age range is given, generally obtained by determining the points on the real calendar that are connected to the points (BP$\pm \sigma$) via the calibration curve.

This procedure is explained in more detail in Mook and Waterbolk (1985) and Pearson (1987). Within the range(s) selected, all dates are equally valid. Computer programs exist to perform such calibrations (Stuiver & Reimer, 1986). Only supplying an age range as such is reasonably correct, but valuable information is lost since the entire range is not equally probable.

Computer programs have been developed to calculate the probability distribution of the calibrated age (for references, see Aitchison *et al*, 1988). We provide a graphic representation of the calibration procedure. Earlier reports on the Groningen program were published elsewhere (Mook, Hasper & van der Plicht, 1987; van der Plicht, Mook & Hasper, in press).

THE GRONINGEN COMPUTER PROGRAM

The computer program developed at the CIO in Groningen runs on an IBM-XT or compatible PC. It is written in Turbo-Pascal version 4.0, (Borland, 1988). The graphics recommended is CGA, but EGA and Hercules are also supported. A mathematical co-processor is optional.

The CAL4.PAS program can perform the calibration interactively. The code asks for the ^{14}C age to be calibrated (in BP), its error (σ BP) and an identification label (user-chosen text). It then reads from an input file (INPUT4.DTA) additional necessary data such as the integration step size along the calendar axis, plot axis scaling option and files containing the calibration data points and spline function data. The file INPUT4.DTA has to be changed only if the standard options are not wanted, eg, calibration data other than Stuiver and Pearson (1986).

A second CAL5.PAS program is a batch mode version of CAL4, which can calibrate several ^{14}C ages automatically in sequence. In this case, the dates to be calibrated (BP), the corresponding errors (σ BP) and text labels are read in from the input file INPUT5.DTA. The graphic results are printed using the CGA-screen dump.

The programs CAL4.PAS and CAL5.PAS, and the necessary data files, are available for distribution. All calibration data as published in the Trondheim Proceedings (Stuiver & Kra, 1986) are provided, with exception of the marine data. More specialized versions (not distributed) can perform tasks such as output routing to a color plotter, calculate "moving averages" (de Jong & Mook, 1980), "wiggle matching" (Pearson, 1986), spline functions along the calibration points and averaging more dates into one calibrated result (van der Plicht, Mook & Hasper, in press).

The calibration points, x, y, correspond to AD/BC, BP. A function y = f (x) has to be determined first which represents the calibration curve. This calculation has to be done only once. We have chosen a "cubic spline function", calculated according to the procedure of Reinsch (1967). With this procedure, a smooth spline is calculated along a set of data points. The calculation is performed iteratively such that

$$\sum_i [(y (x_i) - y_i) / \sigma (y_i)]^2 \le S,$$

where S is the "smoothing parameter." As discussed by Reinsch (1967), recommended values for S depend on the magnitude of the standard deviations $\sigma(Y_i)$. We have chosen $S = 0$, where the routine produces the natural

cubic spline, *ie,* it interpolates the data points. This also means that the error in the calibration data points is not taken into account.

The calculation results in four matrices a_i, b_i, c_i and d_i, which define the function:

$$f(x) = [(d_i \cdot h + c_i) \cdot h + b_i] \cdot h + a_i,$$

where $h = x - x_i$ and $x_i \leq x \leq x_{i+1}$, x_i being the calibration points. The matrices a_i, b_i, c_i and d_i (for each calibration data set, and corresponding to a smoothing parameter $S = 0$) are stored in a file, and called in by the program CAL4.PAS (or CAL5.PAS) following the specifications in INPUT4.DTA (or INPUT5.DTA). Once the spline function $y = f(x)$ is calculated, the actual calibration can be performed, *ie,* the transformation of the ^{14}C age (BP, y-axis) into the real calendar age (cal AD/BC, x-axis) via the function $y = f(x)$.

The ^{14}C age is the result of a statistical physical measurement, and thus approaches a Gaussian probability distribution, which has the form:

$$\exp[-(y - BP)^2/2\sigma_{BP}^2]$$

The probability distribution along the x-axis (the calendar axis) is not a Gaussian, and instead has the form:

$$\exp[-(f(x) - BP)^2/2\sigma_{BP}^2]$$

In order to obtain quantitative results, we need the cumulative probability distribution (which is a summation of the previous expression):

$$P(x) = \int_{-\infty}^{x} \exp\left[-(f(x) - BP)^2 / 2\sigma_{BP}^2\right] dx$$

which is normalized by integrating to $x = \infty$. In fact, the integration boundaries are chosen as those points along the x-axis, corresponding to BP $\pm$ $3\sigma_{BP}$, which corresponds to a probability interval of 99.7%.

The probability function is determined by stepping through the selected area along the x-axis (the calendar axis). This is essential in case there is a wiggle in that part of the calibration curve considered, i e, the function $y = f(x)$ is not monotonous. Mathematically, this means that if a measured date y needs to be transformed into a calendar date x, one needs to use an inverse function $x = g(y)$ allowing the determination of x from y. However, since there are several x's corresponding to the same y, g is not a function in the usual mathematical sense, since a function can only have one value at any point (see Renfrew & Clark, 1974). By integrating along the x-axis, possible mathematical traps are avoided. The probability function $P(x)$ is in fact a measure of the residence time in the calibration curve $f(x)$.

The computer program selects that part of the calibration curve corresponding to the ^{14}C age BP $\pm$ $3\sigma_{BP}$. In the corresponding area along the calendar axis, the function $P(x)$ as defined above is calculated. The integration step size (dx) is usually taken as 4 years (corresponding to 5 steps for the 20-yr intervals).

It is illustrative to consider the case $f(x)$ is a straight line, namely BP = 1950 − AD. Here $P(x)$ corresponds to the well-known probability function
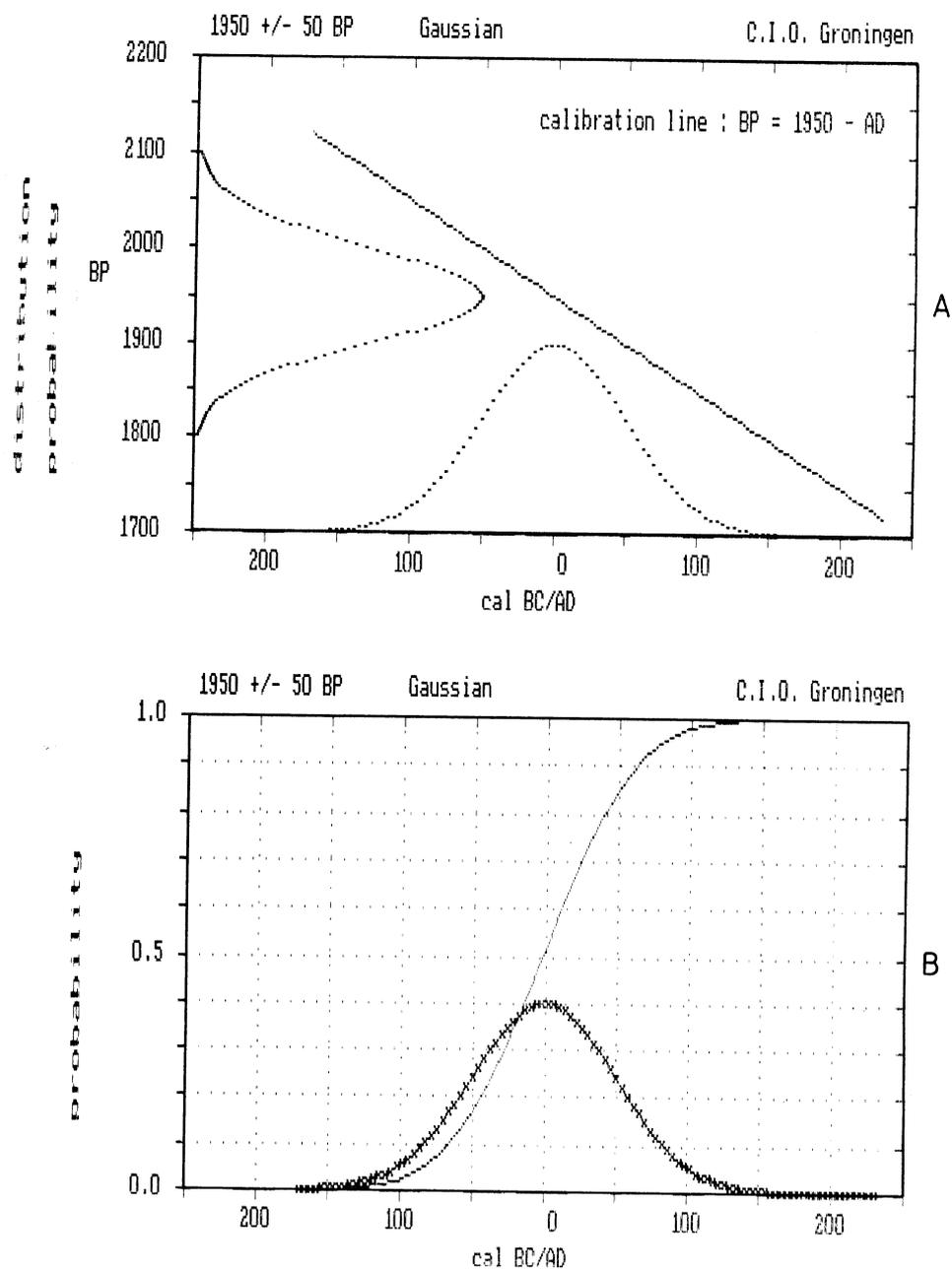
Fig 1. Hypothetical calibration with a straight line as calibration function

for Gaussian distributions (the error function):

$$P(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(- \tfrac{1}{2} x^2)\, dx = \tfrac{1}{2} \left[1 + \text{erf} \left[\frac{1}{\sqrt{2}} \right] \right]$$

This is schematically shown in Figure 1. The calibrated result shows a Gaussian probability distribution, with the same standard deviation as the ^{14}C age. If the straight line is transformed into the wiggly calibration curve, the calibrated result is no longer Gaussian, and the resulting age can no longer be expressed as an average result with an assigned error bar. Examples will be discussed in the next chapter.

EXAMPLES OF PROBABILISTIC CALIBRATIONS

An example of a calibration performed with the Groningen program is shown in Figure 2. The sample dated was GrN–14980 (Satricum B11/3570), the result of which was 2365 ± 45 BP. To determine the calendar age of samples from this period is very difficult. The calibration curve is more or less flat between 750 and 400 cal BC, so that the calibrated dates of samples ca 2450 BP always cover a long range, no matter how precisely the ^{14}C age
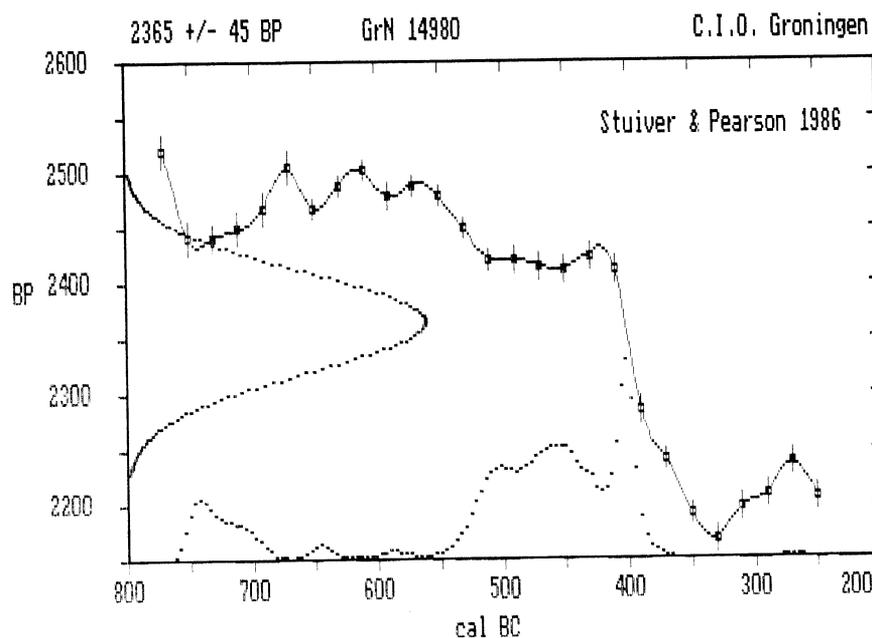
Fig 2. Calibration of a ^{14}C age of 2365 ± 45 BP. The graph shows the calibration curve (spline function) connecting the calibration data points, the Gaussian probability distribution corresponding to the ^{14}C age (dotted line along the y axis) and the calibrated calendar age probability distribution (dotted line along the x axis).

is measured. Beyond 410 cal BC, on the contrary, the calibration curve is very steep so that the calibrated date will have a much narrower peak than the corresponding ^{14}C age.

Therefore, we think this to be a good example to illustrate the calibration procedure. Three curves are shown in Figure 2: the calibration curve (spline function) with the calibration points from Stuiver and Pearson, the Gaussian probability distribution of the ^{14}C age (2365 ± 45 BP) along the y axis, and the resulting calendar age probability distribution along the x axis. The latter is calculated from 250 to 770 BC, corresponding to 2230 (= 2365 − 3 × 45) and 2500 (= 2365 + 3 × 45) BP, respectively. In this probability distribution, features caused by the course of the calibration curve can easily be recognized, such as the wiggles around 750, 650, 575 and 270 cal BC, and the narrow peak at 400 cal BC caused by the steep slope of the calibration curve.

The major problem, once a calibrated probability distribution is obtained, is interpretation. In the previous version of our program, written in PC-MATLAB and discussed in van der Plicht, Mook and Hasper (in press), we suggested two ways to state the calibrated result. First, if the probability distribution has the shape of one bump, the result can be given as the median (which is where the cumulative probability $P(x) = 0.5$ or 50%), with an error range of 68% from $P(x) = 0.16$ to $P(x) = 0.84$. Second, if the calibration yields several (isolated) peaks, the median should *not* be used since this most often corresponds to a calculated date with a low probability, or is even impossible. Further, this approach does not make sense in cases such as shown in Figure 2, the result of calibrating 2365 ± 45 BP. The 16–84% probability interval would, in this case, yield a result of 675 − 415 cal BC, which is obviously unrealistic.

We suggest two ways to interpret the probability distribution (Fig 3). Figure 3A shows the probability distribution along the calendar axis together with the normalized cumulative probability distribution. The peak between 760 and 675 cal BC has a probability of 16%; the peak between 675 and 620, 2%, between 610 and 560, 1%, between 560 and 360, 80%, and the small bump at 275 cal BC, ca 1%. Assigning medians to the isolated bumps at 760–675 and 560–360 cal BC, which would be located at 735 cal BC ($P(x)$ = 8%) and 455 cal BC ($P(x) = 60\%$) is not recommended. (We neglect here the small peaks at 650, 590 and 275 cal BC.) Working with the median (defined as the center of gravity of the distribution) or modus (defined as the maximum) for the peak at 760–675 cal BC is justified, but not for the bump at 560–360 cal BC. In this case, the median is located at 455 cal BC. If we consider the 68% range of the bump, the result could be the range 400–530 cal BC. The modus is located at 400 cal BC, which is just at the boundary of the 68% range. In principle, the modus can even be outside the 68% range. Therefore, in our opinion, using the terms median or modus is not incorrect, but for ^{14}C calibration purposes confusing, and should generally be avoided.

Figure 3B shows an alternative way to interpreting the probability distribution. The same calendar axis probability distribution is shown, although with a different scale (arbitrarily normalized to the maximum of the probability distribution). The two dotted lines correspond to a 68.3 and
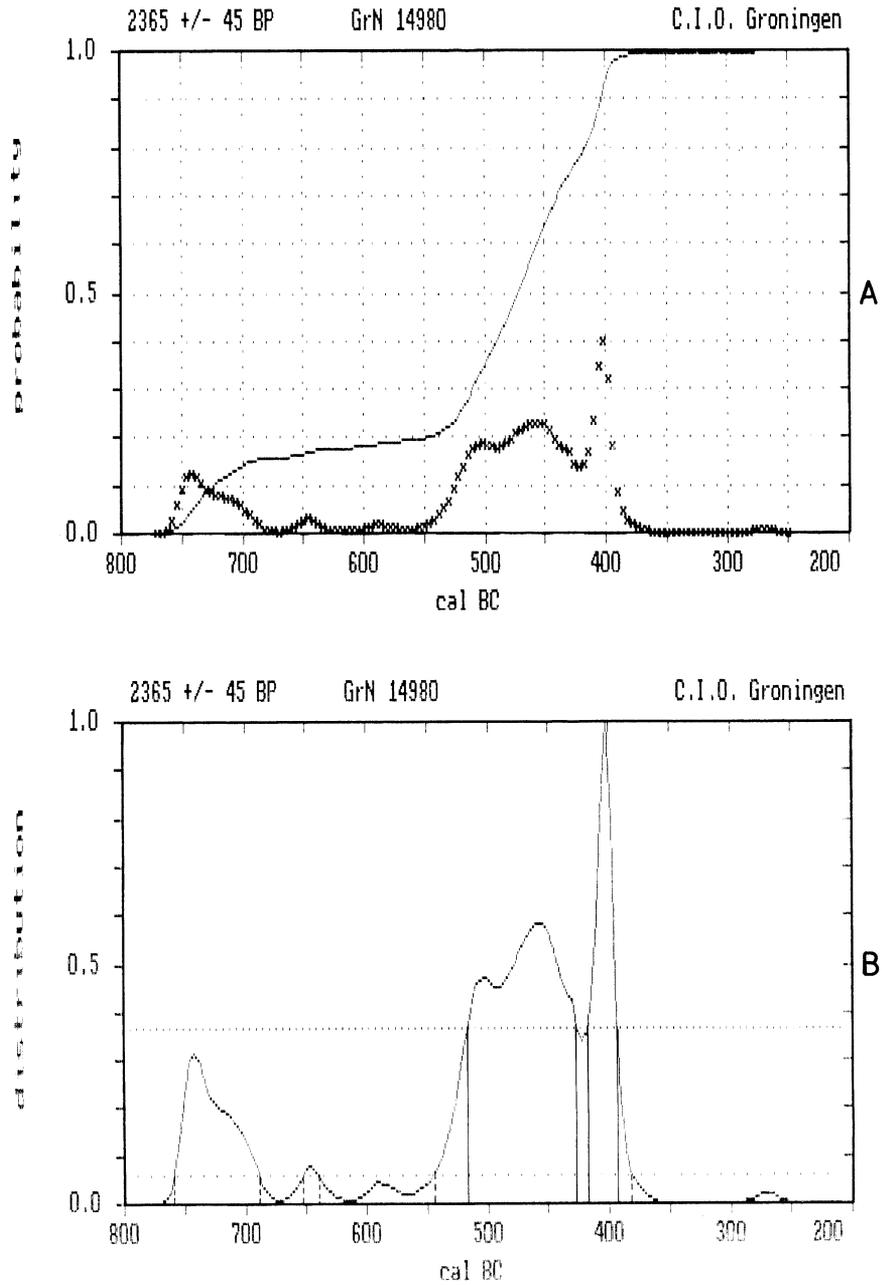
Fig 3. A. Calendar axis probability distribution and the normalized cumulative probability distribution. B. Calendar axis probability distribution with levels corresponding to 1-σ and 2-σ confidence levels (dotted lines).

95.4% confidence level, respectively. This means that the area under the probability distribution curve, taken between the intercepts of the dotted lines and the distribution is 68.3% (resp 95.4%) of the total area. These intercepts (indicated by the vertical lines) can then be taken as the calibrated ranges. For this example, 2365 ± 45BP, the results can be stated as follows: 518–422 and 418–394 cal BC for the 68.3% (1σ) confidence level, and 758–686, 654–638 and 546–378 cal BC for the 95.4% (2σ) level. This analysis follows the procedure of Stuiver and Reimer (1987).

AVERAGING MORE THAN ONE CALIBRATION DATE

To average several calibrated dates, the Gaussians can be taken together to yield a final average which is again a Gaussian distribution with a new smaller standard deviation:

$$< y > = \frac{\sum_i (y_i / \sigma_i^2)}{\sum_i (1 / \sigma_i^2)} \quad \text{and} \quad < \sigma > = \frac{1}{\sqrt{\sum_i (1 / \sigma_i^2)}}$$

This may be followed by the calibration described above. The same procedure is not possible for non-Gaussian distributions such as real calendar age probability distributions.

This procedure cannot always be applied. Calibrated distributions have to be summed for archaeological results, eg, if the contemporaneity of the samples is in question, if real time periods are to be determined, if we want to determine durations of major activities, or if the time width of the sample is larger than the standard deviation of the measurement (Mook, 1983).
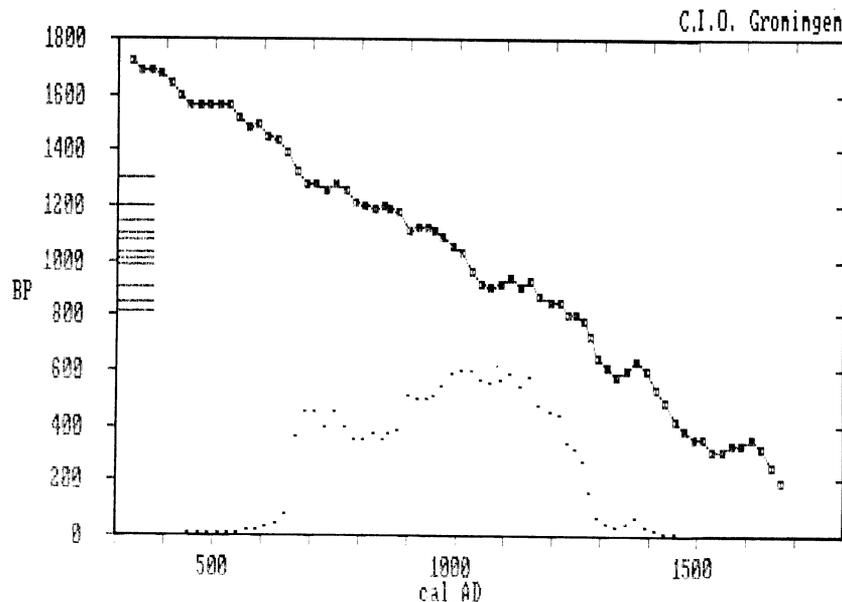
Fig 4. Example of calibrating a series of 11 dates into one cumulative probability distribution

To determine the probability distribution of more than one BP date, we first calibrate each BP date as described above which yields a real calendar age distribution y(x). Results are then normalized so that the area $\int y(x)\,dx = 1$. Only then the individual real age distributions y(x) may be added. Of the total summed function y_{tot} (x), the probability distribution P(x) can again be obtained, again normalized to a total probability of 1 (or 100%). In Figure 4, the ages of 11 samples, from excavations of the St Walburg Church in Groningen with error bars ranging from 40–200 yr, are calibrated and added. The 11 Gaussian probability distributions of the ^{14}C ages are not plotted, but are represented by bars along the y axis. The calibrated probability distribution along the calendar (x) axis can be analyzed as discussed above. We did not pretreat the ^{14}C ages in any way, eg, as outlined by Aitchison *et al* (1989). Thus, the resulting probability distribution, therefore, is rather broad.

OTHER DEVELOPMENTS: MOVING AVERAGE AND WIGGLE MATCHING

Special options are built into the program. First, the calibration curve can be smoothed by the "moving average procedure." For materials representing many years (such as peat or charcoal), a detailed curve has to be smoothed to an extent depending on the time width of the sample. For a 100-yr-old sample measured to ±20-yr precision, a curve averaged over 100 yr should be used rather than a 1-yr curve (Mook, 1983). From the spline calibration function described before, we obtain a new set of calibration points in 1-yr intervals. These calibration points are averaged in groups, with the average value for the calibration function of the group assigned to the average calendar year.

The same procedure is repeated for the next group of calibration points, starting at the next year, etc. In this way, the averaging process moves through the calibration curve. The final result is then a smoothed calibration function (van der Plicht, Mook & Hasper, in press). The 1-yr calibration points obtained in this way are only an approximation. The 20-yr points from Stuiver and Pearson (1986) are measured from 20 tree-ring samples and are, thus, already averaged over these 20-yr intervals. Ideally, we need a 1-yr curve based on actual measurements taken in sufficiently small steps. Such data, however, only exist for isolated time periods (see, eg, de Jong & Mook, 1980).

Second, a special version of the program will be able to wigglematch. This might be successfully applied to small peat samples (Van Geel & Mook, ms).

DISCUSSION AND CONCLUSION

We developed a PC-based computer program for automatic calibration of ^{14}C ages, written in Turbo Pascal (version 4.0). The program uses a spline function generated through the calibration data points, which can be collected from the Radiocarbon Calibration Issue (Stuiver & Kra, 1986), with exception of the marine curve. The calibration data from Stuiver and Pearson are used as a standard, going back to 2490 cal BC, extended by the curve of Pearson *et al* which goes back to 5210 cal BC.
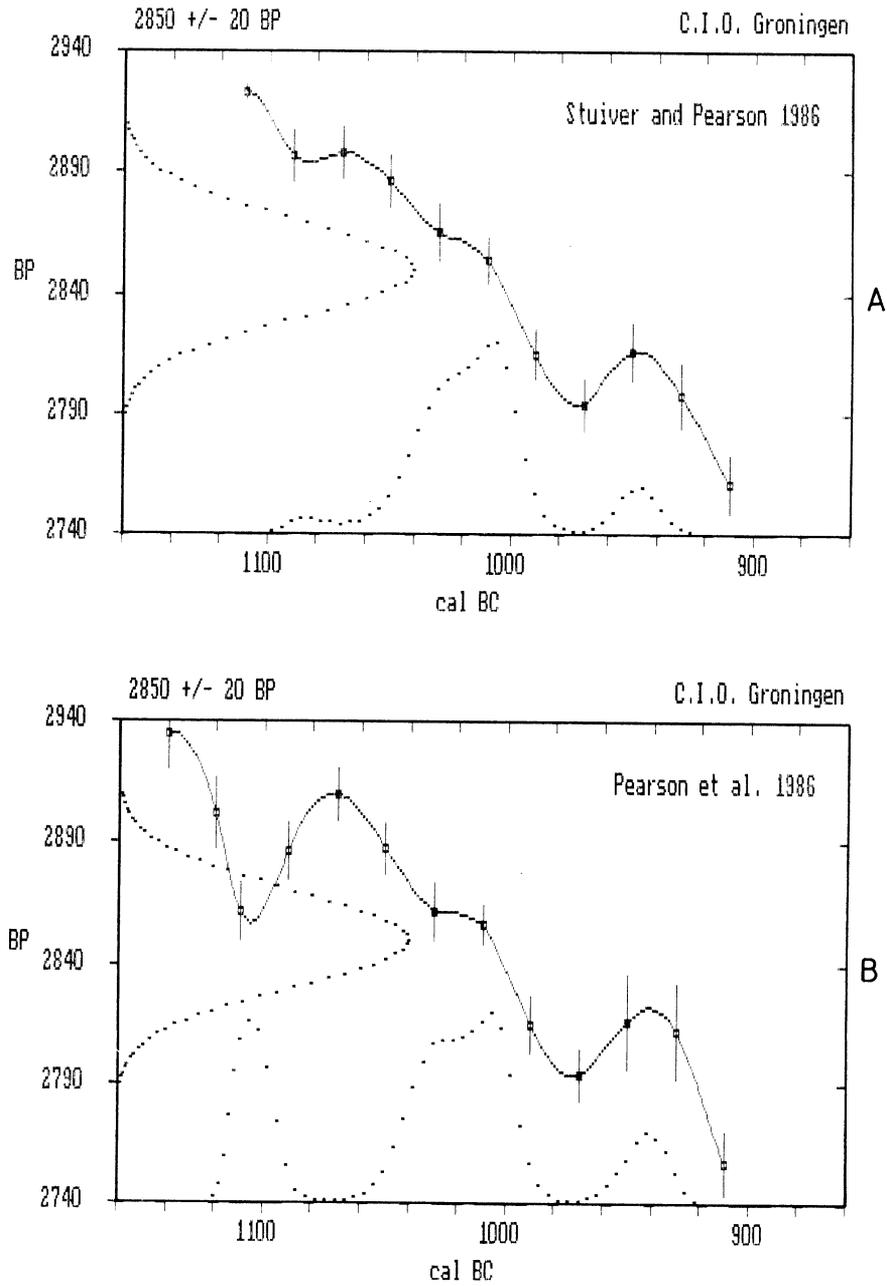
Fig 5. Calibrating a ^{14}C age of 2850 $\pm$ 20 BP using two different calibration curves

The Gaussian probability distribution of the ^{14}C age is taken at the 3σ level. The program automatically selects part of the calibration curve and calculates the calibrated probability distribution along the calendar axis. The program is user friendly (especially for CGA-equipped computers): one only has to enter the date to be calibrated (in BP), the associated error (σ BP) and a text label, either interactively or by entering these data in an input file. Some familiarity with Turbo-Pascal (4.0) is helpful.

The calibrated results are produced mainly in graphic form. Conclusions on calibrated ages or ranges can best be drawn by analyzing the calibrated probability distribution. We offer two options. First, we can use the normalized cumulative probability distribution to obtain quantitative results. From this cumulative probability, the area under the calendar calibrated age distribution, we can directly determine the probability of a calibrated age range or peak (Fig 3A). Unfortunately, the computer graph has to be analyzed manually.

Second, it was agreed during this conference to add to our analysis the program of Stuiver and Reimer (1987), which produces an extra plot of the same probability distribution with indications of the 68.3 (1σ) and 95.4 (2σ) confidence levels. The calibrated age ranges corresponding to these levels can be printed out (Fig 3B).

A special version of the program analyzes an ensemble of dates into one calibrated result, by adding the single probability distributions properly. The ^{14}C ages can also be calculated at the 3σ level. Other special versions can be used for moving average and wiggle matching. The procedure by which the calibrated probability distribution is calculated is, in our opinion, mathematically correct. We should be aware, however, that the obtained distribution is sensitive to certain parameters that determine the shape of the calibration curve. Thus far we have used a spline function curve, obtained according to Reinsch (1967). To calculate the spline function, we have chosen the smoothing parameter at zero, *ie,* the curve connects all calibration data points. We feel this is justified for the ^{14}C calibrations, but spline functions calculated with a nonzero smoothing parameter may be more realistic and may change details of the resulting calibrated probability distributions. The same may be true if we use different calibration curves. Figure 5 shows the same date, 2850 ± 20BP, calibrated with both the recommended curve of Stuiver and Pearson (1986), and that of Pearson *et al* (1986). In the latter, the wiggle at ca 100 BC produces a peak in the calibrated probability distribution with a probability of 25%, whereas in the Stuiver and Pearson curve, the wiggle is smoothed out and the 25% peak disappears almost completely.

REFERENCES

Aitchison, T C, Leese, M, Mook, W G, Otlet, R L, Ottaway, B S, Pazdur, M F, van der Plicht, J, Reimer, P R, Robinson, S W, Scott, M, Stuiver, M, Walker, J and Weninger, B, 1989, Comparison of methods used for the calibration of radiocarbon dates: Radiocarbon, this issue.

Aitchison, T C, Ottaway, B S and Scott, E M, in press, Statistical treatment of groups of related radiocarbon dates, *in* Symposium on archaeology and ^{14}C, 2nd, Proc: PACT.

Borland, 1988, Turbo-Pascal, version 4.0: Borland Internatl, Scotts Valley, California.

Geel, B van and Mook, W G, (ms), High resolution ^{14}C dating of organic deposits using natural atmospheric ^{14}C variations: Ms subm to Nature.

Hassan, F A and Robinson, S W, 1987, High-precision radiocarbon chronometry of ancient Egypt, and comparisons with Nubia, Palestine and Mesopotamia: Antiquity, v 61, p 119–135.

Jong, A F M, de (ms) 1981, Natural ^{14}C variations: PhD dissert, Univ Groningen.

Jong, A F M, de and Mook, W G, 1980, Medium-term atmospheric ^{14}C variations, *in* Stuiver M and Kra R S, eds, Internatl ^{14}C conf, 10th, Proc: Radiocarbon, v 22, no. 2, p 267–272.

Mook, W G, 1983, ^{14}C calibration curves depending on sample time width, *in* Symposium on ^{14}C and archaeology, Proc: PACT, Pub no. 8, p 517–525.

Mook, W G, Hasper, H and Plicht, J, van der, 1987, Background and procedures of ^{14}C calibration, *in* Aurenche, O, Evin, J and Hours, F, eds, 1987, Chronologies in the Near East: CNRS Internatl Symposium, Lyon (France), BAR Internatl ser 379(i).

Mook, W G and Waterbolk, H T, 1985, Handbook for archaeologists no. 3: Radiocarbon dating, European Sci Found.

Pearson, G W, 1986, Precise calendrical dating of known growth-period samples using a "curve fitting" technique, *in* Stuiver, M and Kra, R S, eds, Internatl ^{14}C conf, 12th, Proc: Radiocarbon, v 28, no. 2B, p 839–862.

—— 1987, How to cope with calibration: Antiquity, v 61, p 98–103.

Pearson, G W, Pilcher, J R, Baillie, M G L, Corbett, D M and Qua, F, 1986, High-precision ^{14}C measurement of Irish oaks to show the natural ^{14}C variations from AD 1848–5210 BC, *in* Stuiver, M and Kra R S, eds, Internatl ^{14}C conf, 12th, Proc: Radiocarbon, v 28, no. 2B, p 911–934.

Pearson, G W and Stuiver, M, 1986, High-precision calibration of the radiocarbon time scale, 500–2500 BC, *in* Stuiver M and Kra, R S, eds, Internatl ^{14}C conf, 12th, Proc: Radiocarbon, v 28, no. 2B, p 839–862.

Plicht, J, van der, Mook, W G and Hasper, H, in press, Automatic calibration of radiocarbon dates, *in* Symposium on archaeology and ^{14}C, Groningen, 2nd, Proc: PACT.

Reinsch, C H, 1967, Smoothing by spline functions: Num Mathematik, v 10, p 177–183.

Renfrew, C and Clark, R M, 1974, Problems of the radiocarbon calendar and its calibration: Archaeometry, v 16, p 5–18.

Stuiver, M and Kra, R S, eds, 1986, Calibration issue, Internatl ^{14}C conf, 12th, Proc: Radiocarbon, v 28, no. 2B, p 805–1030.

Stuiver, M and Pearson, G W, 1986, High-precision calibration of the radiocarbon time scale, AD 1950–500 BC, *in* Stuiver, M and Kra, R S, eds, Internatl ^{14}C conf, 12th, Proc: Radiocarbon, v 28, no. 2B, p 805–838.

Stuiver, M and Reimer, P J, 1986, A computer program for radiocarbon age calibration, *in* Stuiver, M and Kra, R S, eds, Internatl ^{14}C conf, 12th, Proc: Radiocarbon, v 28, no. 2B, p 1022–1030.

—— 1987, User's guide to the program Calib and Display Rev 2.1: Quaternary Isotope Lab, Univ Washington.

Suess, H E, 1970, The three causes of secular C14 fluctuations, their amplitudes and time constants, *in* Olsson, I U, ed, Nobel Symposium, 12th, Proc: Stockholm, Almqvist & Wiksell, p 595–605.

Vries, H, de, 1958, Variation in concentration of radiocarbon with time and location on earth: Koninkl Nederlandse Akad Wetensch Proc, v 61, ser B, p 1–9.

Willis, E H, Tauber, H and Münnich, K O, 1960, Variations in the atmospheric radiocarbon concentration over the past 1300 years: Am Jour Sci Radiocarbon Suppl, v 2, p 1–4.