[RADIOCARBON, VOL 31, NO. 3, 1989, P 824–832]

IMPROVEMENT OF THE PROCEDURE FOR PROBABILISTIC CALIBRATION OF RADIOCARBON DATES

MIECZYSŁAW F PAZDUR and DANUTA J MICHCZYŃSKA
Radiocarbon Laboratory, Institute of Physics, Silesian Technical University, Krzywoustego 2, PL-44-100 Gliwice, Poland

ABSTRACT. A set of computer procedures for probabilistic calibration of ^{14}C dates was developed at the Gliwice Radiocarbon Laboratory for the IBM PC compatible microcomputer. The program comprises three main options: 1) calibration of a single ^{14}C date, 2) calibration of a set of arbitrary dates, 3) calibration of a set of related dates. Results of calibration are presented in the form of graphs and numeric data, including tables of selected quantiles and interquantile ranges of resulting probability distribution of cal age. In this paper, we present the aims of the program, with a short description of its structure, show examples of working with output data in terms of expected archaeological application, and consider the possibility of standardization of calibration procedures.

INTRODUCTION

A set of computer procedures for probabilistic calibration of single ^{14}C dates and sets of arbitrary or related dates was developed for the IBM PC/XT microcomputer with the Hercules graphic card. The theory of probabilistic calibration was presented during the Symposium "Archaeology and ^{14}C" in Groningen, 1987 (Michczyńska, Pazdur & Walanus, in press). The calibration procedure consists in transforming the initial Gaussian probability distribution on the calendric time scale to the appropriate part of the calibration curve. Several changes were made in the program after discussions during the Groningen meeting and with archaeologists.

AIMS OF THE PROGRAM

Any set of computer procedures for calibration of ^{14}C dates should be designed following several important requirements, which may be listed as follows:

1) the algorithm of calibration should be mathematically correct and complete,

2) the program should be user-oriented as much as possible,

3) the presentation of output data should not impose any limitations on further interpretations (within obvious limits of statistical validity),

4) for archaeological applications, the program should take into account the most specific tasks, which include calibration of a single ^{14}C date, calibration of a set of arbitrary dates to show their distribution on the calendric time scale, calibration of a set of repeated dates, representing the same level, event, or sample, estimation of the duration of a specific phase or culture, comparisons between different cultures/phases, testing their contemporaneity or estimating probable time lags, ordered in the calendric scale according to certain criteria. The first requirements seem to be obvious, while others may be disputable.

Before going into the details of the calibration procedure, we must clarify our terminology. We must distinguish clearly between the terms "calibration," "standardization" and "interpretation of calibration output." Calibration itself means transforming the initial conventional ^{14}C date (= its probability distribution) onto the calendric time scale. The calibration procedure should provide an adequate representation of transformed probability distribution in terms of graphs and point and/or interval estimates. The decision on graphic presentation of calibration results presents no problems as only two plots are in common use, *ie,* the probability density function and cumulative probability. The choice between point and interval estimates is more complicated, and because of the great variability of shapes of resulting probability distributions of the calibrated age, it seems that both forms of presentation should be incorporated in a standard version of the calibration procedure. There is some question on combining them in a reasonable way to obtain the most universal starting point for further interpretation of the data.

TABLE 1

Problems in standardization of calibration procedures

Level	Subject	Adopted solution	Q*
Input data	Calibration curve	Stuiver & Pearson (1986) Pearson & Stuiver (1986) Pearson *et al* (1986)	HI, O, N
	Form of probability distribution of conventional ^{14}C date	Gaussian	HI, O, N
	Truncation level	3σ or 4σ	HI, O, N
Methods	Smoothing	Y/N: if Y, how?	I, NO, U
	Error of cal curve	Y/N; if Y, how?	I, NO, N
Output data	Graphs	Probability density function Cumulative probability	I, O, U
	Numeric data Point estimate Ranges	Both	NI, NO, U

Q* = quantification
HI = highly important; I = important; NI = not important; O = obvious; NO = not obvious; N = necessary; U = unnecessary

Standardization should be considered at three levels: 1) input data, 2) methods of treatment and 3) output data. The guidelines for standardization of computer calibration procedures are presented in Table 1. The choice at the first level is obvious, as international agreement was reached for the curves of Stuiver and Pearson (1986) and Pearson and Stuiver (1986) for calibrating ^{14}C dates younger than 2500 BC. It seems reasonable to use the curve of Pearson *et al* (1986) as a natural extension of the two former curves. Standardization of methods used for treatment of calibration data is more difficult to achieve, as different methods of smoothing and including the errors of the cal curve may be introduced at this stage.

Our decision for taking the calibration curves as they were published, *ie,* in the form of piece-line functions, ignoring their errors, is supported by two arguments. First, for medium-accuracy ^{14}C dates ($\sigma \geq 50$ yr), the effect of smoothing and the contribution of error on the calibration curve seem to be negligible; even a small offset equal to several years may lead to serious misinterpretations of high-accuracy dates or mean values of series of repeated datings.

PROCEDURES

The first option of the program, to calibrate a single ^{14}C date, remains unchanged. The output contains graphic information (plots of probability density and cumulative probability functions) and numeric data (locations of maxima of probability, range/ranges of cal age, selected values of quantiles and selected interquantile ranges).

In option 2, calibrating a set of arbitrary dates, a simple bar plot is displayed on the screen, showing intervals of cal age which are cut from the calibration curve by a band of width equal to [D – MN*σ, D + MN*σ], where D and σ denote the conventional ^{14}C date and its standard deviation, and MN may be equal to 1, 2, 3 or 4. In the second version of this option, we can obtain a "two-dimensional" plot showing location of the median, as well as interquartile ranges and 95% confidence intervals of the calibrated age of all dates included in the calibration. The list of all dates with appropriate data is available on the printer.

In option 3, calibrating a set of related dates, the output contains plots of cumulative probability and probability density functions with tables of selected quantiles and interquantile ranges. If applicable or desired, the weighted mean of all conventional dates is calculated and is next calibrated as for a single date. In all three options, both graphs and numeric data may be displayed in either BP or AD/BC notation. For options 2 and 3, the number of dates included in the calibration is not limited by the program. Data sets containing ~ 50 dates can be analyzed in a reasonable time.

EXAMPLES

To illustrate how the requirements listed above are achieved by the program and to discuss some general problems of interpretation of calibration output, we will consider two examples taken from a data set prepared for the project on the comparison of calibration procedures (*cf* Aitchison *et al,* 1989).

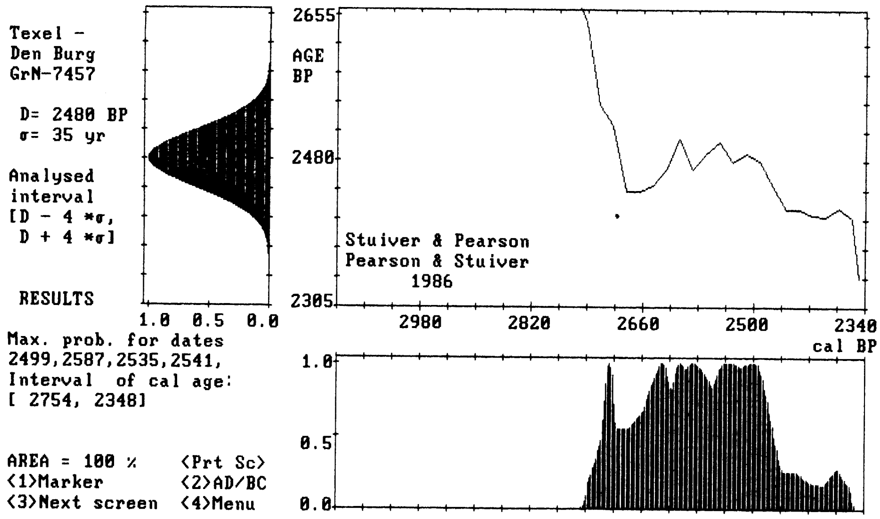

Fig 1. Calibration of a single date – copy of the first screen of cal output

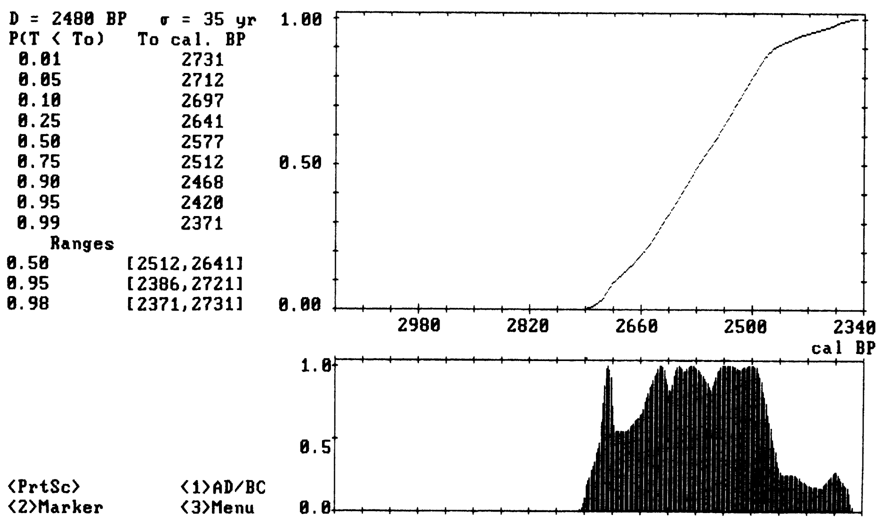

Fig 2. Calibration of a single date – copy of the second screen of cal output

https://doi.org/10.1017/S0033822200012443 Published online by Cambridge University Press

The result of calibrating a single date of sample Texel-Den Burg (GrN–7457: 2480 ± 35 BP) is shown in Figures 1 and 2. The first screen of cal output (Fig 1) contains relevant introductory information about input and output data and shows the appropriate section of the calibration curve. More detailed results are shown on the second screen (Fig 2). The shape of the resulting probability density function of cal age may be regarded as approximately rectangular (uniform) in the range of cal ages from ca 2480 to 2720 cal BP. In such a case, the median (equal to 2576 cal BP ≈ 2580 cal BP) should be taken as the midpoint value of cal age. The choice of appropriate interval of cal age which may be used to characterize the uncertainty of the calibrated date is not obvious, and because such a choice should be, in principle, based on specific tasks and the decision of the user, the output data do not impose any limitations.

Three ranges of cal age, corresponding to confidence levels 0.50 (interquartile range), 0.95 ("2σ" interval), and 0.98 ("3σ" interval) are given explicitly; others may be easily derived either from a listed table of quantiles or from a plot of cumulative probability. The "1σ" confidence interval, eg ranges from 2490 – 2670 cal BP. Summarizing, date 2480 ± 35 BP may be represented in the calendric time scale by combination of point estimate (median) and the three most widely used confidence intervals, *ie*, as 2580^{+60}_{-70} cal BP (interquartile range), or 2580±90 cal BP (68% or "1σ" interval), or 2580^{+140}_{-200} cal BP (95% or "2σ" interval).

The second example illustrates the calibration of two sets of ^{14}C dates of Michelsberg II (10 dates from various sites) and Michelsberg III (9 dates from the same site) cultures (Lanting & Mook, 1977, p 60–61). Figure 3

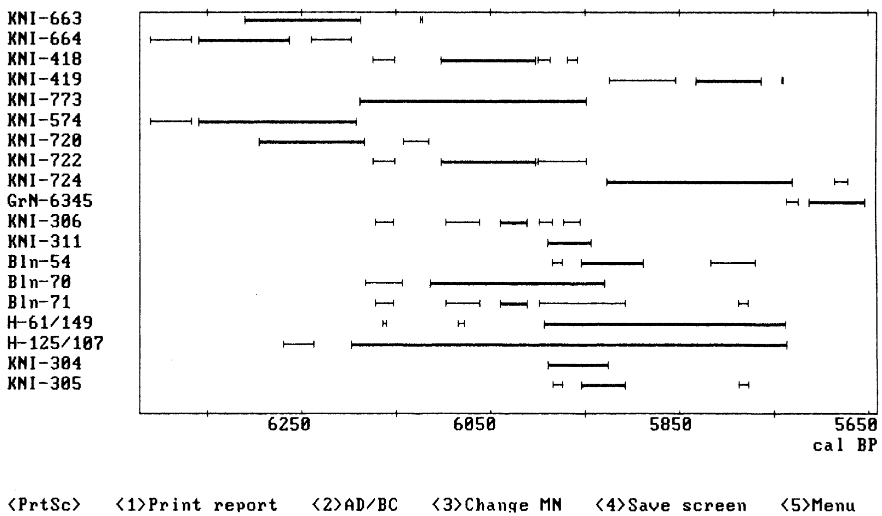

Fig 3. Michelsberg II (top) and Michelsberg III (bottom) cultures. Bar plot obtained in the "1σ–cutout" calibration – option 2A of calibration of arbitrary dates. Bold lines denote intervals of maximum probability.

https://doi.org/10.1017/S0033822200012443 Published online by Cambridge University Press

shows the distribution of ^{14}C dates on the calendric time scale (obtained with option 2 of the calibration program). Bars on the plot show intervals of calendric age which are cut from the calibration curve by a band of width [D+σ, D–σ]; bold bars indicate intervals of maximum probability. "Calibrated" dates of the Michelsberg II culture almost uniformly span an interval of more than seven centuries, while those of Michelsberg III are grouped ca 6000 cal BP.

This presentation is useful for a rough evaluation of calibrated age values, but it seems too complicated for long data series. A more useful picture is obtained using the second version of option 2, in the form of a plot (Fig 4). This shows all relevant information resulting from the calibration of

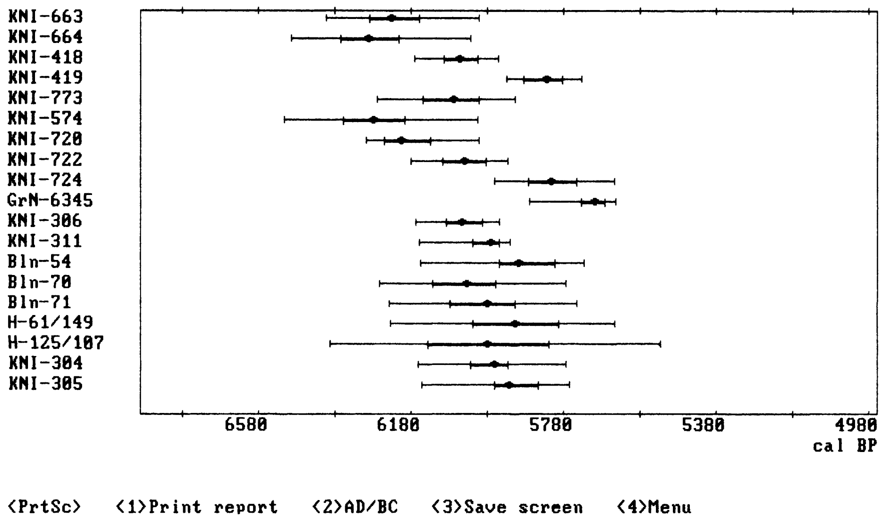

Fig 4. Michelsberg II (top) and Michelsberg III (bottom) cultures. Bar plot obtained in "two-dimensional" calibration – option 2B of calibration of arbitrary dates showing location of median (dot), interquartile range (bold) and 95% confidence interval (regular line).

each date in a data set in a simplified graphic form (including median, interquartile ranges and 95% confidence intervals). Each plot may be complemented with a table containing laboratory numbers, conventional ^{14}C dates and appropriate ranges of cal ages. Table 2 shows the calibration results in this version of the procedure.

An estimate of duration of cultures can be obtained using option 3. Figure 5 shows an example of the calibration output for Michelsberg II. The resulting probability density function of cal age can be characterized as two approximately uniform parts; the first, ca 5700–5970 cal BP, the second, 5970–6270 cal BP, with a long tail extended towards older ages. This form of composite probability distribution of cal ages of 10 samples from the Michelsberg II culture may be, in fact, expected from bar plots shown in

https://doi.org/10.1017/S0033822200012443 Published online by Cambridge University Press

TABLE 2

List of dates included in calibration
Data set: Michelsberg II and III

No.	Sample no.	Age ^{14}C conv BP	Median cal BC	Confidence intervals cal BC 50%	95%
1	KNI–663	5440± 85	–4283	[–4341, –4211]	[–4455, –4054]
2	KNI–664	5490± 95	–4345	[–4417, –4266]	[–4549, –4076]
3	KNI–418	5270± 40	–4106	[–4148, –4058]	[–4223, –4005]
4	KNI–419	5080± 50	–3878	[–3938, –3836]	[–3983, –3784]
5	KNI–773	5280± 85	–4123	[–4202, –4055]	[–4321, –3958]
6	KNI–574	5480±105	–4331	[–4410, –4249]	[–4567, –4057]
7	KNI–720	5400± 60	–4260	[–4304, –4183]	[–4352, –4056]
8	KNI–722	5250± 60	–4095	[–4151, –4035]	[–4235, –3981]
9	KNI–724	5050± 85	–3863	[–3925, –3797]	[–4015, –3697]
10	GrN–6345	4965± 40	–3750	[–3787, –3723]	[–3922, –3696]
11	KNI–306	5260± 40	–4099	[–4142, –4044]	[–4221, –4001]
12	KNI–311	5210± 40	–4024	[–4071, –4002]	[–4212, –3973]
13	Bln–54	5140± 80	–3950	[–4002, –3854]	[–4207, –3779]
14	Bln–70	5240±100	–4088	[–4177, –4009]	[–4315, –3825]
15	Bln–71	5200±100	–4033	[–4131, –3958]	[–4293, –3796]
16	H–61/149	5140±130	–3958	[–4071, –3846]	[–4285, –3698]
17	H–125/107	5200±200	–4032	[–4190, –3869]	[–4448, –3577]
18	KNI–304	5190± 60	–4014	[–4076, –3978]	[–4214, –3827]
19	KNI–305	5160± 60	–3976	[–4015, –3900]	[–4205, –3816]

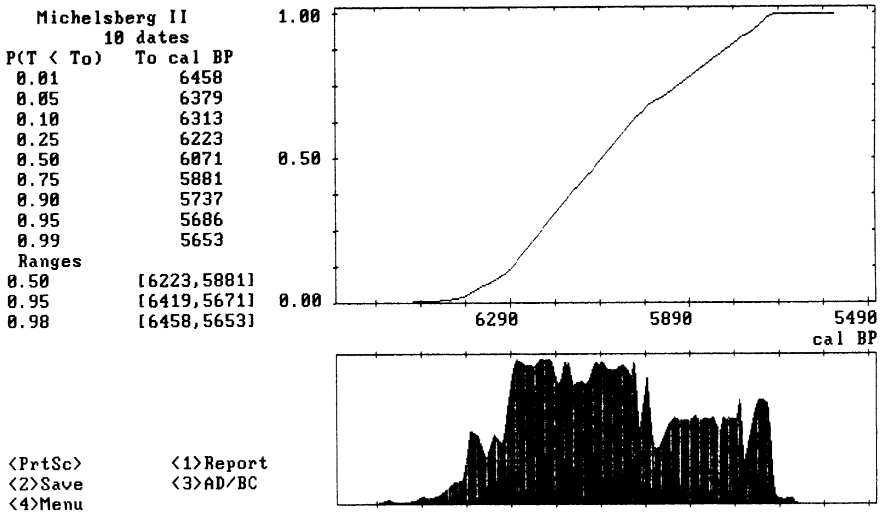

Fig 5. Michelsberg II culture. A composite probability distribution of calendric ages of 10 dates from option 3 of calibration of a set of related dates.

https://doi.org/10.1017/S0033822200012443 Published online by Cambridge University Press

TABLE 3

Calibrated ^{14}C chronology of Michelsberg II and III cultures

Culture	50% confidence level			95% confidence level		
	Cal BP	Cal BC	Duration (yr)	Cal BP	Cal BC	Duration (yr)
Michelsberg II	6220–5880	4270–3930	340	6410–5670	4460–3720	740
Michelsberg III	6070–5910	4120–3960	160	6250–5710	4300–3760	540
Michelsberg II & III	6150–5900	4200–3950	250	6390–5670	4440–3720	720

Figures 3 and 4. The composite probability distribution of cal ages of 9 dates of Michelsberg III has quite a different shape; it has a prominent peak at ca 5960 cal BP and 3 other peaks of approximately the same height, occurring at 5830, 6060 and 6120 cal BP.

The duration of cultures can be estimated in either 50% (interquartiles) or 95% confidence intervals. The results are summarized in Table 3. We concluded that both cultures overlap and decline at approximately the same time, although Michelsberg II originates 150 yr earlier than Michelsberg III. Exactly the same conclusions concerning the relationship between these cultures are drawn from 50% and 95% interval estimates, despite significant differences in the duration of the intervals. The results of joint treatment of dates belonging to both cultures, shown also in Table 3, yield exactly the same 95% estimates as those obtained for the Michelsberg II culture.

CONCLUSIONS

Because of the specific wiggle shape of recent high-precision calibration curves, the procedures for calibrating ^{14}C dates should be based on the transformation of probability distributions of conventional ages into a calendric time scale. The resulting probability distribution in most cases significantly differs from the initial Gaussian distribution and general rules for simple presentations of calibration output cannot be formulated. Moreover, because of the same reason, the concepts that are widely used and familiar to non-experts in statistics and probability (eg, mean value, median, 1σ confidence interval) begin to lose their seemingly unshakable credibility. The probability distributions derived from our computer programs graphically present and describe both point and interval estimates. Based on the shape of the probability density function, the user can choose the most appropriate set of parameters to characterize the calibration result.

Finally, most archaeological applications involve series of ^{14}C dates. We have taken this into account by including various possibilities for the treatment of groups of dates; a 2- or 3-stage procedure is recommended. At the first introductory stage, the group of dates should use option 2 to obtain a bar plot ("1σ" intervals or "two-dimensional" distribution of calendric ages) showing the relationship between calibrated ages of individual dates. At the second stage, the composite probability distribution of calendric ages, of all the dates in the group can be obtained using option 3. At the third stage, calibration of the mean value can be performed in the same way as for a single ^{14}C date.

ACKNOWLEDGMENTS

This study was supported by a grant from Central Research Project CPBP 01. 06. The authors also wish to acknowledge the stimulating correspondence with B Weninger.

REFERENCES

Aitchison, T C, Leese, M, Michczyńska, D J, Mook, W G, Otlet, R L, Ottaway, B S, Pazdur, M F, van der Plicht, J, Reimer, P R, Robinson, S W, Scott, E M, Stuiver, M and Weninger, B, 1989, A comparison of methods used for the calibration of radiocarbon dates, *in* Stuiver, M and Kra, R S, eds, Internatl ^{14}C conf, 13th, Proc: Radiocarbon, this issue.

Lanting, J N and Mook, W G, 1977, The pre- and protohistory of the Netherlands in terms of radiocarbon dates: Groningen, 247 p

Michczyńska, D J, Pazdur, M F and Walanus, A, in press, Bayesian approach to probabilistic calibration of radiocarbon dates, *in* Archaeology and ^{14}C symposium, 2nd, Proc: PACT.

Pearson, G W, Pilcher, J R, Baillie, M G L, Corbet, D M and Qua, F, 1986, High-precision ^{14}C measurements of Irish oaks to show the natural ^{14}C variations from AD 1840–5210 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.

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.