USING THE BAYESIAN METHOD TO STUDY THE PRECISION OF DATING BY WIGGLE-MATCHING

TOMASZ GOSLAR and WIESŁAW MĄDRY

Institute of Physics, Silesian Technical University, Krzywoustego 2, PL-44-100 Gliwice, Poland

ABSTRACT. The "wiggle-matching" technique has been widely used for the absolute dating of a series of radiocarbon-dated samples connected in one floating chronology. This is done by calculations of SS statistics (the mean-square distance of ^{14}C ages of samples from the calibration curve) calculated for any assumed calendar age of the floating chronology. In the standard procedure the confidence intervals of true calendar age are derived from the width of the SS minimum, using the critical values of the chi-square distribution. This, however, seems oversimplified. Another approach is an extension of the Bayesian algorithm for calibration of single ^{14}C dates. Here, we describe in detail the Bayesian procedure and discuss its advantages compared to the SS minimization method. Our calculations show that for given errors of ^{14}C measurements, precision of dating the series is related to the shape of the SS curve around its minimum, rather than to the absolute value of SSmin. In some cases, dating precision may be improved more efficiently by extending the time span covered by the series rather than by improving the precision of the ^{14}C measurements. The application of the Bayesian method enabled us to delimit the age of the floating varve chronology from the sediments of Lake Gościąż with distinctly better accuracy than was previously reported using the SS curve alone.

INTRODUCTION

Because of variations in atmospheric radiocarbon concentrations in the past, the result of ^{14}C dating— a ^{14}C age—usually differs from the "true" calendar age of sample. To estimate the calendar age, one must use the calibration curve, which shows the dependence of ^{14}C *versus* calendar ages. The most precise part of the calibration curve relies on high-precision ^{14}C dates of tree rings and covers the last *ca.* 11 ka. A special issue of *Radiocarbon* (1993) contains the calibration data obtained on bidecadal or decadal samples of wood from Europe and North America.

Variations in the past ^{14}C concentration can influence the precision of dating by the ^{14}C method. For a single sample, the precision of a calibrated date is usually worse than the error of ^{14}C age. This is especially evident for those periods where the calibration curve reveals large wiggles, and where the same ^{14}C age corresponds to a few calendar-age values. On the other hand, those wiggles can improve the precision age determination when we are dealing with a series of samples and know exactly the differences of calendar age between them. Such a series may consist of wood samples included in one tree-ring sequence, but for some reason not dated absolutely by dendrochronology. A similar case is the layers from an annually laminated lake sediment.

The Wiggle-Matching Technique

The wiggle-matching technique (or "curve-fitting") has been used for dating of tree-ring sequences for 20 yr (*e.g.*, Beer *et al.* 1979; Kruse *et al.* 1980; Linick, Suess and Becker 1985) and later on, has been also applied to sequences of annually laminated lake sediments (*e.g.*, Hajdas, Bonani and Goslar 1995; Goslar *et al.* 1995). The wiggle-matching technique was described in detail by Pearson (1986). With this method, one looks for the age of the series (represented by the age of the oldest sample, T_s) that gives the best fit of obtained dates to the calibration curve. The quality of the fit is expressed by the mean-square difference (normalized to the dating error) between ^{14}C ages of samples and the ages derived from the calibration curve. In practice, this difference (SS) is plotted *versus* the age of series, and the one with the lowest SS value is regarded as the most probable value of T_s.

Pearson (1986) discussed the question of uncertainty of such an estimate, and noticed that if T_s is true, the statistics n·SS (where n is the number of samples) has a χ^2 distribution, and argued that the confidence intervals of χ^2/n (for any given probability P) can be used directly to determine confidence intervals of T_s. He proposed using the interception of the upper confidence limit of χ^2/n with the SS curve (Fig. 1*a*). This procedure, however, sometimes leads to erroneous results if we want to derive the probability distribution of T_s. An obvious example is the case of samples (*e.g.*, n=3) descending from the period of a wide ^{14}C age plateau (Fig. 1b). In this case, the SS curve has a broad minimum, and the minimum SS value is *ca.* 1. If the interceptions of critical χ^2/n values for P=0.95 (7.8/3=2.6) and for P=0.68 (3.5/3=1.2) trace the 0.95 and 0.68 confidence intervals for T_s, the probability that T_s lies within the residual interval (out of the 0.68 but still within the 0.95 interval) is 0.27. The first challenge is determining what fraction of that probability corresponds to the left and what to the right part of the residual interval. To make matters worse, the residual interval is very narrow, which makes probability density of T_s higher in the residual than in that of 0.68. This means that any T_s is more probable at the edges than at the bottom of broad minimum!

One must agree that the confidence limits of χ^2/n *concern only the values of SS* if the age T_s is real, but apparently they are not transferable to the *domain of the age* itself.

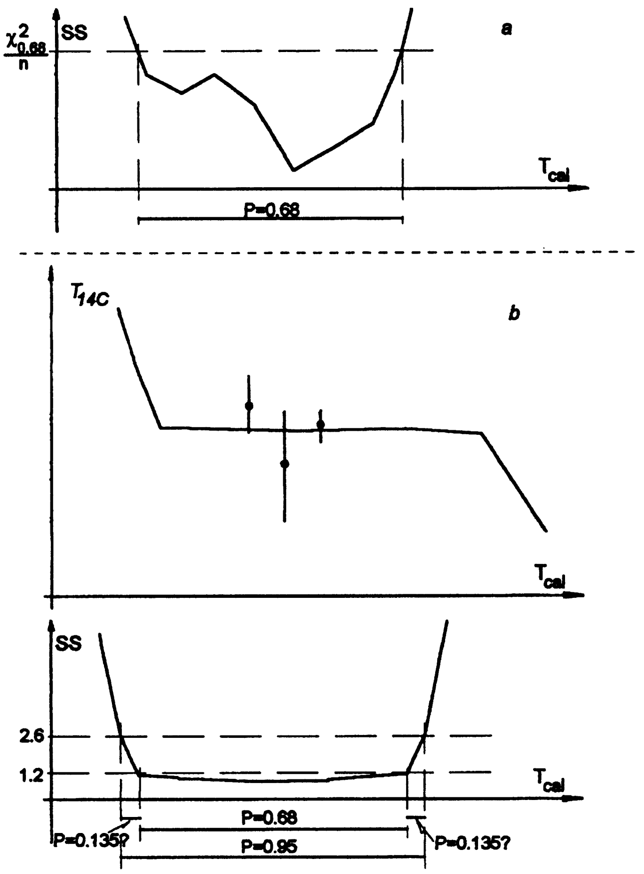

Fig. 1. a. Illustration of the method for deriving the confidence intervals of the age determined by the standard wiggle-matching technique (Pearson 1986). b. Illustration of the case where the standard technique produces incorrect results. Details are given in the text.

The Bayesian Approach

The Bayesian approach enables us to calculate directly the probability distribution of the age of series T_s. It is the natural extension of the algorithm for calibration of single ^{14}C dates (Michczyńska

https://doi.org/10.1017/S0033822200018415 Published online by Cambridge University Press

et al. 1990; van der Plicht *et al.* 1990; van der Plicht 1993), and it also has been included in some advanced programs for calibration the series of ^{14}C dates (*e.g.*, Bronk Ramsey 1995).

The crucial assumption for that approach is that we have no *a priori* information about the age of the series. In other words, every age is *a priori* equally probable. For simplicity, we focus only on the interval covered by the calibration curve and assume that the calibration curve is known exactly. We assume also that our ^{14}C dates (denoted below as $T_1, T_2, \dots T_n$) are free of systematic error and that the one-sigma standard dating errors are not underestimated. The method is as follows:

1. We assume some value for T_s in the interval covered by the calibration curve.
2. If T_s is known, the true values of ^{14}C ages of all samples in the series can be read from the calibration curve. The probabilities that individual samples will be ^{14}C dated to $T_1, T_2, \dots T_n$ can then be derived from Gaussian distributions around true dates. The dispersions of those distributions are equal to the one-sigma standard errors of dating.
3. As the results of measurements are independent of one another, probability for obtaining the whole set of ^{14}C dates, $P(T_1, \dots, T_n) = P(T_1) \cdot \dots \cdot P(T_n)$ is a product of appropriate probabilities for individual samples.
4. We repeat steps 1 through 3, assuming other values of T_s, and reconstruct $P(T_1, \dots, T_n)$ as a function of T_s for the whole interval covered by the calibration curve. After normalizing to 1, this function represents the conditional probability for obtaining the set of dates $T_1, \dots, T_n$ if the calendar age of the series is T_s. We denote that conditional probability as $P(T_1, \dots, T_n|T_s)$.
5. We may now reconstruct the probability of T_s using the Bayesian formula,

$$P(T_s|T_1, \dots, T_n) \sim P(T_1, \dots, T_n|T_s) \cdot \mathbf{P}(T_s), \qquad (1)$$

where $\mathbf{P}(T_s)$ denotes the *a priori* probability distribution of T_s. As all the values of T_s are *a priori* equally probable, $\mathbf{P}(T_s)$ = const, we just obtain

$$P(T_s|T_1, \dots, T_n) \sim P(T_1, \dots, T_n|T_s)\,. \qquad (2)$$

In the numerical algorithm, the spaces between assumed values of T_s are discrete. For simplicity we use the T_s values equally spaced ($\Delta T_s = 1$) in the whole interval.

Comparison with the Standard Bayesian Calibration of Single Dates

The method presented here is a natural extension of the Bayesian algorithm for calibration of ^{14}C dates (Michczyńska, Pazdur and Walanus 1990; Stuiver and Reimer 1993). The key to this extension is that our whole set of ^{14}C dates is compared to the set of Gaussian distributions derived from the calibration curve, using known intervals between calendar ages and only one parameter—the absolute age of the whole series. For a single date, our method gives exactly the same result as the standard calibration procedure (Fig. 2). Of course, smoothing the calibration curve distinctly affects the final result.

One must remember that, regardless if it is smoothed or not, the calibration curve does not represent the exact relationship between calendar and ^{14}C ages. Because of the uncertainty of the calibration data, the dispersions of Gaussian distributions of true ^{14}C dates (see step 2 in the description of the matching procedure) should be modified according to the formula

$$\Delta T_{i,mod} = \sqrt{\Delta T_i + \Delta T_{cc,i}^2}, \qquad (3)$$

where ΔT_i is the error of ^{14}C dating of i-th sample, and $\Delta T_{cc,i}$ is the uncertainty of ^{14}C age in the appropriate fragment of calibration curve. In practice, the values of $\Delta T_{cc,i}$ are calculated from 1-σ

https://doi.org/10.1017/S0033822200018415 Published online by Cambridge University Press

errors of calibration ^{14}C data, by interpolation between appropriate data points. The effect of calibration curve uncertainty is usually minor (Fig. 2), and is significant only if the errors of ^{14}C dates of the matched series (ΔT_i) are very small.

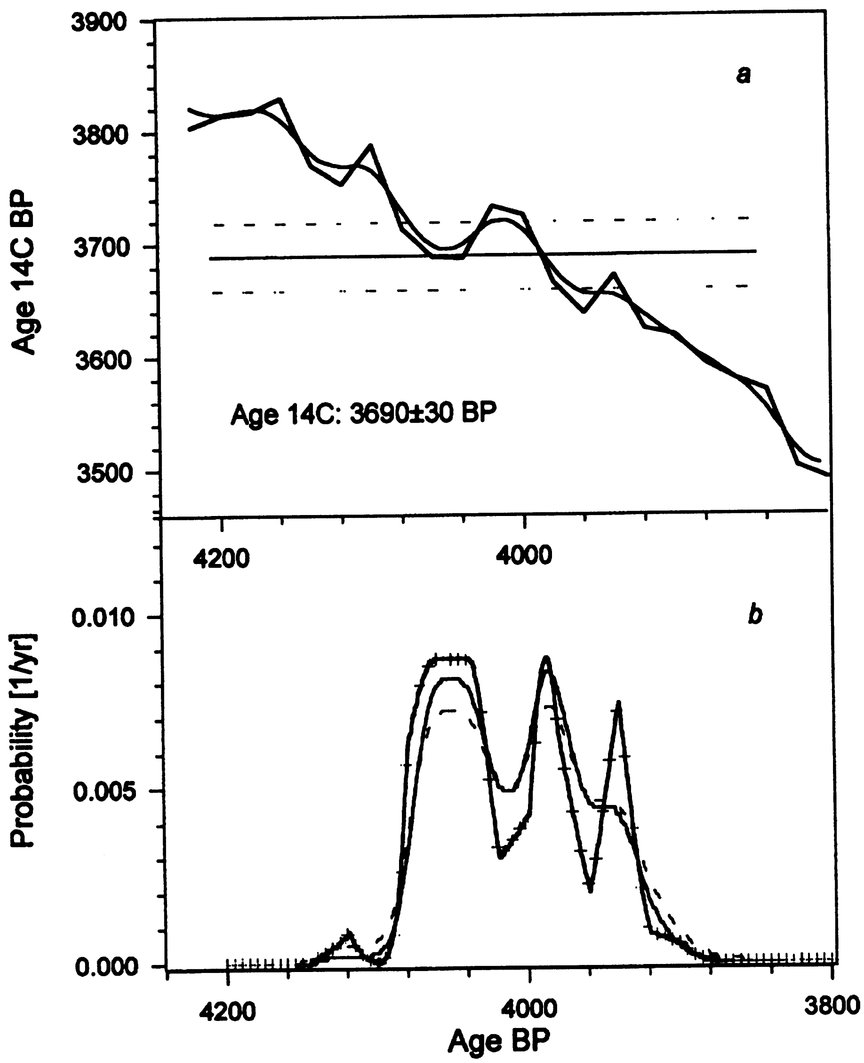

Fig. 2. Comparison of probability distributions given by the procedure use in this work (——) and by the standard method of calibration of individual ^{14}C dates (+). Also illustrated are the effects of smoothing the calibration curve (——), and uncertainty of a smoothed calibration curve (- - -) on the probability distribution.

Comparison with the Standard Wiggle-Matching Procedure

Figure 3 shows a comparison of the SS(T_s) curves and P distributions. The shapes of corresponding SS and P curves closely resemble each other. In fact, the dependence between them is almost straightforward, as

$$P(T_s|T_1, \ldots, T_n) = \prod_{i=1}^{n} P(T_i|T_s) \sim \prod_{i=1}^{n} \frac{1}{\Delta T_i} \exp\left(-\frac{(T_i - T_{c,i})^2}{2\Delta T_i^2}\right) , \quad (4)$$

while

$$SS = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{(T_i - T_{c,i})^2}{\Delta T_i^2} \tag{5}$$

where ΔT_i denotes the error of i-th ^{14}C measurement, and $T_{c,i}$ is the "true" ^{14}C date of i-th sample, derived from the calibration curve, and dependent on the value of T_s.

Hence, if all the errors in the series are similar, we get the approximate relation

$$P(T_s|T_1, \ldots, T_n) \sim \exp \sum_{i=1}^{n} -\frac{(T_i - T_{c,i})^2}{2\Delta T_i^2} = \exp\left(-\frac{n \cdot SS}{2}\right). \tag{6}$$

However, even if all the errors are the same, the values of probability cannot be directly read from the SS curve, since the distribution P must be normalized to 1. The probability

$$P(T_s|T_1, \ldots, T_n) = \frac{\exp\left(-\frac{n \cdot SS}{2}\right)\Delta T_s}{\sum \exp\left(-\frac{n \cdot SS}{2}\right)\Delta T_s} \tag{7}$$

depends thus on the shape of *the whole* curve, and not only on the absolute value of SS for given T_s.

The question of absolute values of SS is illustrated in Figure 3, where the match of three date sets has been compared. The ^{14}C dates in all sets are the same, but the errors in the second and third sets are 2 and 4 times larger, respectively. The best precision of dating the series occurs for the series with the smallest errors, despite the fact that in that case, the highest SS values are obtained. Obviously, the width of probability distribution reflects the width of the SS minimum rather than its absolute level. Of course, if the minimum SS value is very low or very high, the quoted dating errors may be regarded as unrealistic. The question of whether we should estimate the errors from the particular set of few samples, or rather rely on the results of interlaboratory comparisons of many other dates, is a delicate problem, and its discussion is beyond the scope of this paper.

Examples and Discussion

The most uncertain calibrated ^{14}C ages are obtained for periods when the calibration curve has long plateaus. As illustrated in Figure 4, uncertainty of the age of the series may be reduced by the improvement of ^{14}C dating precision, but in certain cases, dating of supplementary samples, extending the time interval covered by the series, may be more effective. It is clear that the 0.68 and 0.95 confidence intervals suggested by the interceptions of SS curve at the critical values of χ^2 (Fig. 4*c*) are quite unrealistic.

As an example of application, we used the Bayesian approach to date absolutely the floating varve chronology of the Lake Gościąż sediments (Fig. 5) and to determine the calendar age of the Younger Dryas/Preboreal (YD/PB) boundary, reconstructed in those sediments. Relying on the width of SS curves, that age was previously determined at 11,440 ± 120 yr BP (Goslar *et al.* 1995). Here, we used the Bayesian procedure for two series of AMS ^{14}C dates, obtained on different series of samples in two different institutions (CFR, Gif-sur-Yvette, France; Goslar *et al.* 1995, and ETH, Zürich; Hajdas, Bonani and Goslar 1995). We also smoothed the calibration curve (Stuiver and Pearson 1993; Pearson and Stuiver 1993; Pearson, Becker and Qua 1993) by the spline function (Reinsch 1967). The fragment of the calibration curve before 7200 BP has been shifted by 41 yr, according to the recently published revision based on matching of the German oak dendroscales (Björck *et al.* 1996).

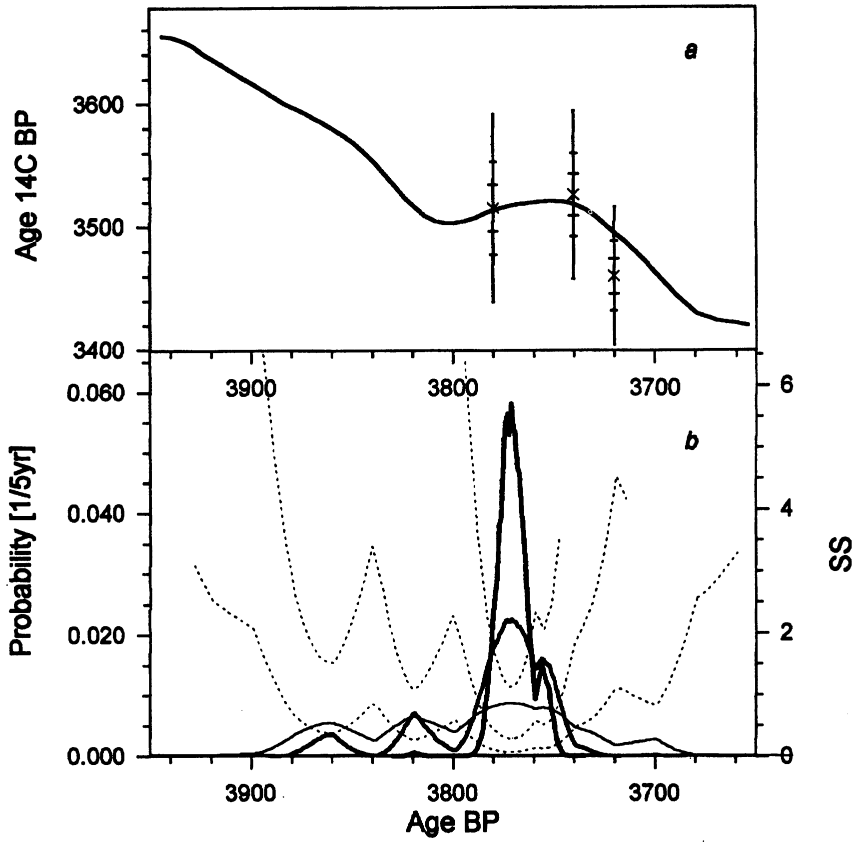

Fig. 3. Comparison of the results of wiggle-matching and Bayesian procedure obtained for different errors of ^{14}C dates. a. smoothed calibration curve and data points with three different values of error. The ratio of errors is 1:2:4. b. Probability distributions (——) and SS curves (- - - -) obtained for different errors of ^{14}C dates.

To collect enough material for AMS dating, the macrofossils for a few samples were collected from sections comprising as many as 150 varves. The uncertainty of sample position with respect to varve chronology was then significant when compared to errors of AMS dating. This has been taken into account by the further modification of dispersions of Gaussian distributions (step no. 2 of the matching procedure)

$$\Delta T_{i,mod} = \sqrt{\Delta T_i^2 + \Delta T_{cc,i}^2 + \Delta \mathrm{T}_i^2} \tag{8}$$

where $\Delta \mathrm{T}_i$ denotes the error of sample position (*i.e.*, the half-thickness of section). Such an approach seems fully correct only if the calibration curve is a straight line with the slope of 45°. This is not exactly the case and the modification should depend also on the shape of the calibration curve. The development of more adequate procedure will be the subject of further study.

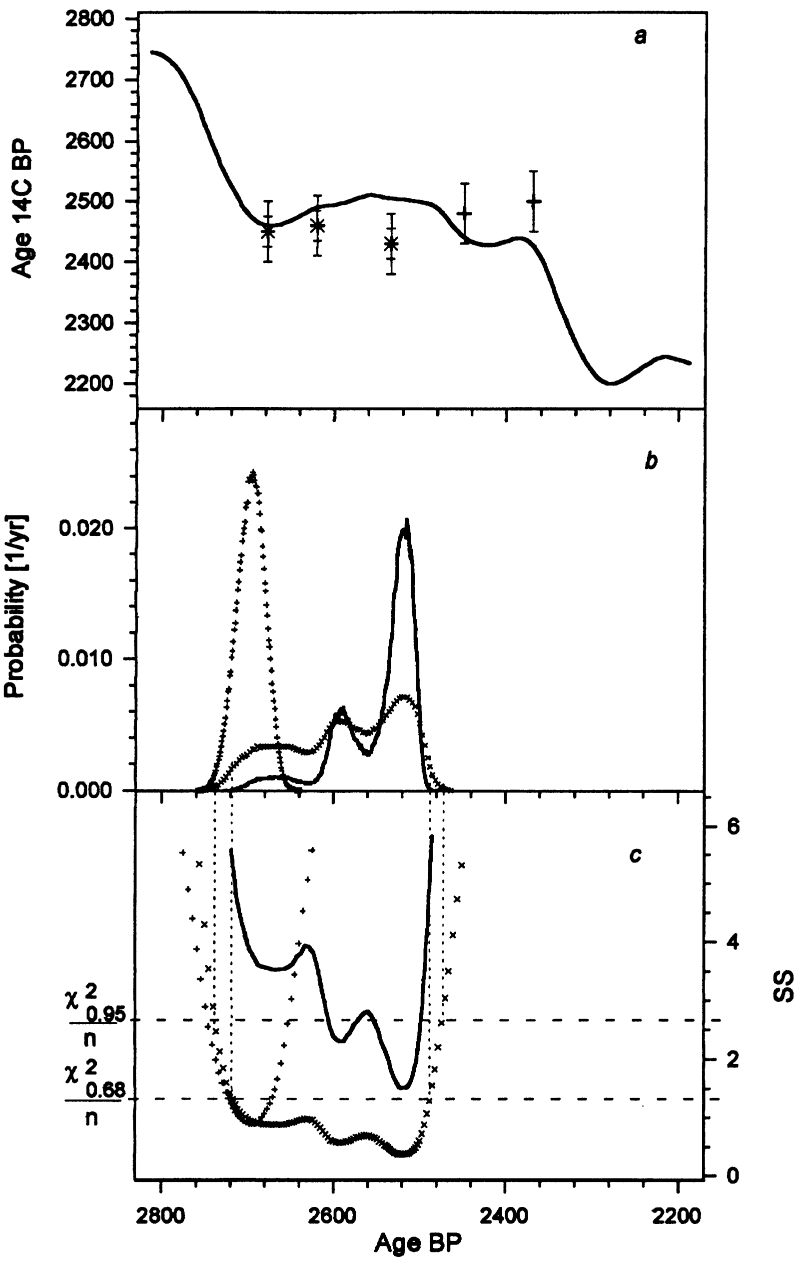

Fig. 4. Comparison of wiggle-matching and Bayesian procedure results obtained for different series of dates from the plateau of the ^{14}C calibration curve. a. smoothed calibration curve and data points. The series of 3 dates (×) was extended to 5 dates (+). b. Probability distributions of the age of the series of 3 dates (×), 3 dates with the errors reduced by a factor of two (——), and 5 dates (+). c. SS curve obtained for the different series of dates. The method of interceptions SS curve at the critical chi-square values is applied to the set of 3 dates.

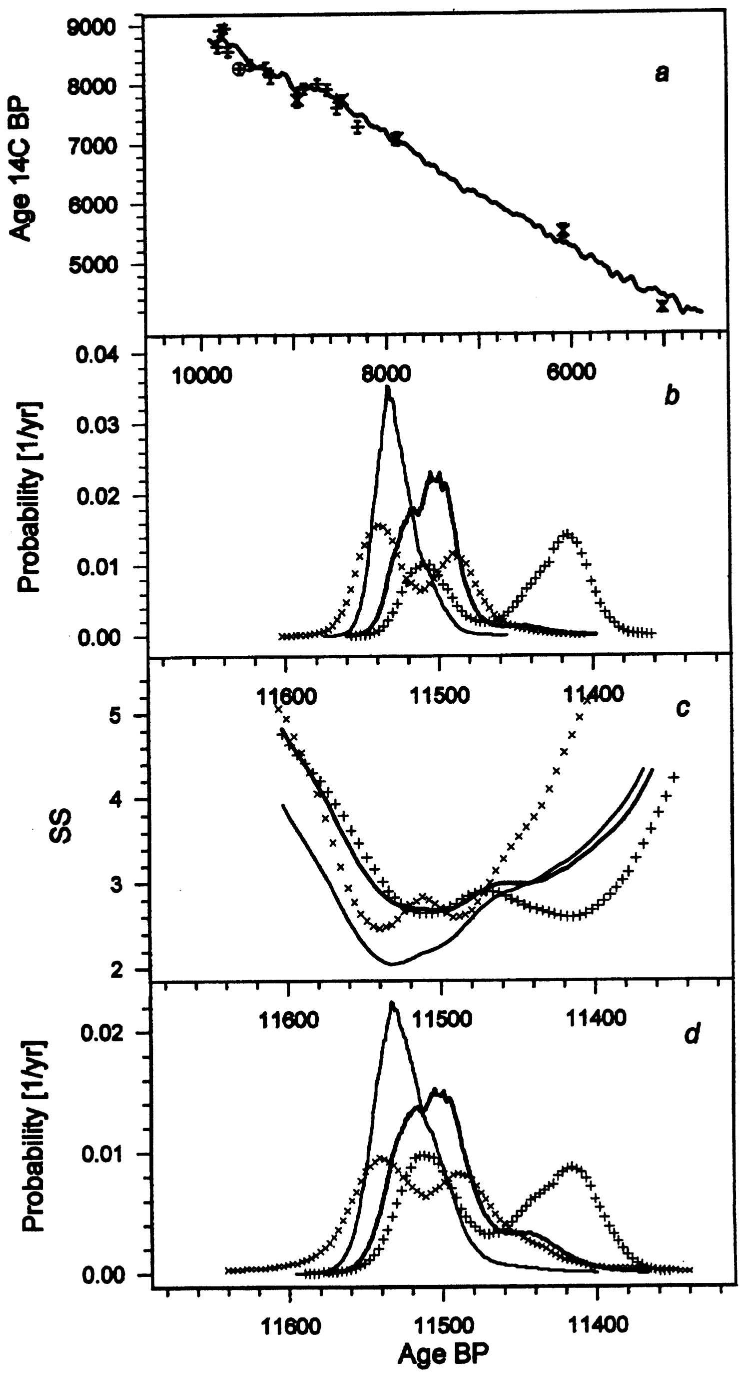

Fig. 5. Results of wiggle-matching and Bayesian procedures, obtained for the series of AMS ^{14}C dates from Lake Gościąż. a. smoothed calibration curve and data points; b. probability distributions obtained for the quoted laboratory errors of ^{14}C measurements; c. the minima of SS curves; d. probability distributions obtained after adjustment of the errors to the values expected from the SS curves (see text for details). Series are denoted by symbols: (×) dates obtained in ETH, Zürich (Hajdas, Bonani and Goslar 1995); (+) dates obtained in CFR, Gif-sur-Yvette, France; (——) the joint series; (——) the joint series, with one critical date rejected. The critical date is indicated in Fig. 5a by a circle.

Probability distributions of age of the YD/PB boundary, determined for both separate series and for the joint series, agree quite well with one another (Fig. 5b). Uncertainty of the match is distinctly lower for the joint series than for individual ones. Judging by the values of the absolute minima of the SS curves (Fig. 5c), one could suspect that the errors of ^{14}C dates were quoted too low. This could reflect some overestimation of laboratory precision of ^{14}C measurement, but could also be an effect of some (not controlled and possibly variable) delays between growth of particular macrofossil and its deposition in the sediments. Another reason might be accidental contamination of individual samples by modern carbon. Goslar *et al.* (1995) indicated two such samples in the whole set of Lake Gościąż ^{14}C dates. One of these samples lies within the range of calibration curve (Fig. 5a). Rejection of this critical date makes the SS minimum distinctly lower (Fig. 5c), though apparently still too high.

If we agree that the errors of ^{14}C dates are too low indeed, we may only multiply all the errors in the series by the constant factor. Here, such a value of factor was chosen, which normalizes the SS minimum to its expected value. Distributions of the YD/PB age, obtained after such procedure (Fig. 5d), are slightly wider than for the original errors, and, for safety, these results seem preferable. The rejection of critical date produces shift of the distribution, by 15–20 yr towards the older age. This shift seems large when compared with the width of distributions obtained using the original errors of ^{14}C dates (Fig. 5b). However, in the more realistic case, with the errors enlarged to normalize the SS minimum, the shift of match appears insignificant.

The uncertainty of the match (± 25 yr) is much lower than the average error of the individual ^{14}C date in the matched series (± 85 yr). The reduction of uncertainty is a well-known effect of multiple measurement, where the individual errors of independent results partly cancel one another. Adopting the primitive formula

$$\Delta\bar{x} = \Delta\frac{x}{\sqrt{n}} \tag{9}$$

where $\bar{x}$ is the mean value of x, and n is the number of results, we can roughly estimate the precision of match to *ca.* 20 yr. One could expect that the uncertainty is ultimately limited by the precision of ^{14}C dates constituting the calibration curve. However, wiggle-matching of several ^{14}C dates uses several calibration data, and the effect of multiple measurement concerns both data sets. Of course, an increase in the number of measurements does not reduce match uncertainty when ^{14}C dates are affected by systematic error. This problem, however, is beyond the scope of this paper.

Acknowledgments

The authors wish to thank Dr. Adam Walanus for critical reading of the manuscript and many valuable comments. This work was sponsored by the Polish Committee of Scientific Research through grant no. 6 P04E 027 10 to T.G.

References

Beer, J., Giertz, V., Möll, M., Oeschger, H., Riesan, T. and Storhm, C. 1979 The contribution of the Swiss lake dwellings to the calibration of radiocarbon dates. *In* Berger, R. and Suess, H. E., eds., *Proceedings of the 9th International ^{14}C Conference*. University of California Press: 567–584.

Björck, S., Kromer, B., Johnsen, S., Bennike, O., Hammarlund, D., Lemdahl, G., Possnert, G., Rasmussen, T. L., Wohlfarth, B., Hammer, C. U. and Spurk, M. 1996 Synchronised terrestrial-atmospheric deglacial records around the North Atlantic. *Science* 274: 1155–1160.

Bronk Ramsey, C. 1995 Radiocarbon calibration and analysis of stratigraphy: The OxCal program. *In* Cook, G. T., Harkness, D. D., Miller, B. F. and Scott, E. M., eds., Proceedings of the 15th International ^{14}C Conference. *Radiocarbon* 37(2): 425–430.

Goslar, T., Arnold, M., Bard, E., Kuc, T., Pazdur, M. F., Ralska-Jasiewiczowa, M., Rózanski, K., Tisnerat, N., Walanus, A., Wicik, B. and Wieckowski, K. 1995 High concentration of atmospheric ^{14}C during the Younger Dryas cold episode. *Nature* 377: 414–417.

Hajdas, I., Bonani, G. and Goslar, T. 1995 Radiocarbon dating of Holocene part of the Gościąż floating varve chronology. *Radiocarbon* 37(1): 71–74.

Kruse, H. H., Linick, T. W., Suess, H. E. and Becker, B. 1980 Computer-matched radiocarbon dates of floating tree-ring series. *Radiocarbon* 22(2): 260–266.

Linick, T. W., Suess, H. E. and Becker, B. 1985 La Jolla measurements of radiocarbon in South German oak tree-ring chronologies. *Radiocarbon* 27(1): 20–32.

Michczyńska, D. J., Pazdur, M. F. and Walanus, A. 1990 Bayesian approach to probabilistic calibration of radiocarbon ages. *In* Mook, W. G. and Waterbolk, H. T., eds., Proceedings of the Second International Symposium, ^{14}C and Archaeology, Groningen 1987. *PACT* 29: 69–79.

Pearson, G. W. 1986 Precise calendrical dating of known growth-period samples using a "curve-fitting" technique. *In* Stuiver, M. and Kra, R., eds., Proceedings of the 12th International ^{14}C Conference. *Radiocarbon* 28(2A): 292–299.

Pearson, G. W., Becker, B. and Qua, F. 1993 High-precision ^{14}C measurement of German and Irish oaks to show the natural ^{14}C variations from 7890 to 5000 BC. *In* Stuiver, M., Long, A., and Kra, R. S., eds., Calibration 1993. *Radiocarbon* 35(1): 93–104.

Pearson, G. W. and Stuiver, M. 1993 High-precision bidecadal calibration of the radiocarbon time scale, 500–2500 BC. *In* Stuiver, M., Long, A. and Kra, R. S., eds., Calibration 1993. *Radiocarbon* 35(1): 25–33.

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

Stuiver, M., Long, A. and Kra, R. S., eds. 1993 Calibration 1993. *Radiocarbon* 35(1): 1–244.

Stuiver, M. and Pearson, G. W. 1993 High-precision bidecadal calibration of the radiocarbon time scale, AD 1950–500 BC and 2500–6000 BC. *In* Stuiver, M., Long, A. and Kra, R. S., eds., Calibration 1993. *Radiocarbon* 35(1): 1–23.

Stuiver, M. and Reimer, P. J. 1993 Extended ^{14}C data base and revised CALIB 3.0 ^{14}C age calibration program. *In* Stuiver, M., Long, A. and Kra, R. S., eds., Calibration 1993. *Radiocarbon* 35(1): 215–230.

van der Plicht, J. 1993 The Groningen radiocarbon calibration program. *In* Stuiver, M., Long, A. and Kra, R. S., eds., Calibration 1993. *Radiocarbon* 35(1): 231–237.

van der Plicht, J., Mook, W. G. and Hasper, H. 1990 Automatic calibration of radiocarbon ages. *In* Mook, W. G. and Waterbolk, H. T., eds., Proceedings of the Second International Symposium, ^{14}C and Archaeology, Groningen 1987. *PACT* 29: 81–94.