

IS IT POSSIBLE TO FIND A GOOD POINT ESTIMATE OF A CALIBRATED RADIOCARBON DATE?

Adam Michczyński

Silesian University of Technology, Institute of Physics, Radiocarbon Laboratory, GADAM Centre of Excellence, Bolesława Krzywoustego 2, 44-100 Gliwice, Poland. Email: Adam.Michczynski@polsl.pl.

ABSTRACT. The result from probabilistic calibration of a radiocarbon date is given in the form of a probability density function. Consequently, reporting a 68% or 95% confidence interval has become a commonly accepted practice. However, many users of ^{14}C dates still try to present the results of calibration as a single point. This manner of presentation is often applied during the construction of age-depth models due to its convenience and simplicity. In this paper, the author tests whether it is possible to find a good point estimate of a calibrated ^{14}C date. The idea of the tests is to compare, using computer simulation, the true value of the calendar age with the age calculated based on the probabilistic calibration of the ^{14}C date and the method of finding the point estimate. The test is carried out for the following point estimates: mode, median, average, the central point of the confidence intervals, and the local mode inside the confidence intervals. The results show that none of these may be considered as a good estimate.

INTRODUCTION

The procedure for a probabilistic calibration of radiocarbon dates was described about 20 yr ago (Aitchison et al. 1989; Stuiver and Reimer 1989) and is now widely known and universally applied. The result of probabilistic calibration of a ^{14}C date is given in the form of the probability distribution (actually, the probability density function); therefore, reporting 68% and 95% confidence intervals has become a commonly accepted practice. However, many users of ^{14}C dates still try to present the results of calibration as a single point, e.g. the center of the 95% confidence interval or date corresponding to the maximum of the probability density function. This manner of presentation has been applied especially often during construction of age-depth models because of its convenience and simplicity. Some users of ^{14}C dates still apply the intercept method, which is not probabilistic.

The problem of point estimation for calibrated ^{14}C dates has been discussed on various occasions, but it is still an object of discussion. Telford et al. (2004) have analyzed how different estimates of the calibrated ^{14}C date behave as a consequence of small changes of the ^{14}C date. They assume that the value of a good point estimate should change only a little with a small change of the ^{14}C date. They conclude that there is no point estimate that can adequately describe the calibrated ^{14}C date, and wherever possible the full probability distribution should be used; however, when the point estimate must be used, they propose to apply the mean or median.

This paper tests another method to see whether it is possible to find a good point estimate of a calibrated ^{14}C date. The idea of the test is to compare the true value of the calendar age of a sample (i.e. the value we want to estimate) with the age calculated based on probabilistic calibration of the ^{14}C date and the method of finding the point estimate. Because we usually do not know the true value of the calendar age of the sample, this comparison must be done using computer simulation.

METHODS

The idea behind the computer simulation, which allows us to compare the true value of the calendar age with the estimated value of the calendar age, is shown schematically in Figure 1. It may be described by the following steps:

1. We assume the value of the true calendar age of the sample, X_T .

2. The value X_T is transformed by the calibration curve in order to find the ^{14}C date X_R corresponding with X_T .
3. We assume a value of uncertainty (error) of the ^{14}C date, σ .
4. The ^{14}C date $X_R \pm \sigma$ described by the Gaussian probability distribution is calibrated, and we obtain the probability distribution of the calibrated ^{14}C date.
5. We estimate a few statistical parameters of the distribution, which may be useful to estimate the calendar age of the sample, and calculate the difference between the estimated value of the parameter and the true calendar age.

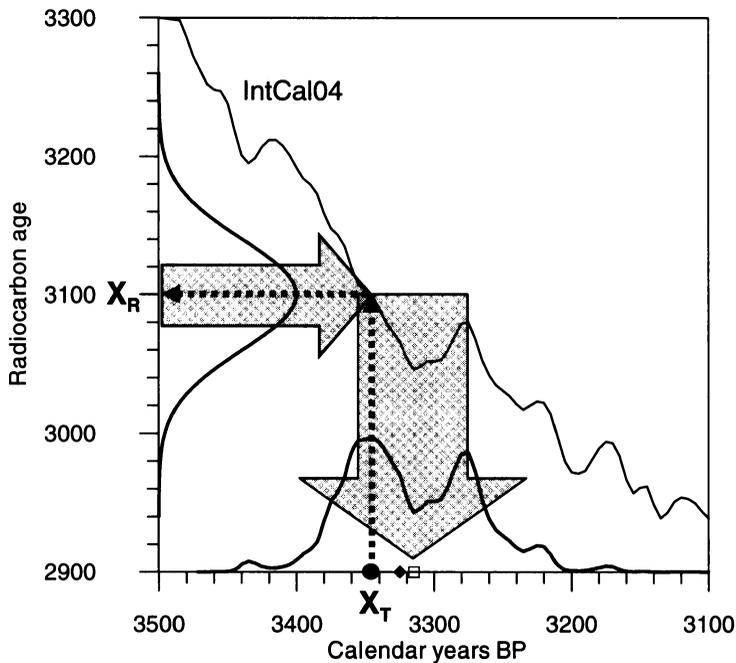


Figure 1 Schematic presentation for the computer simulation (detailed description in text). The true calendar age is marked with a gray circle. A black triangle marks the mode of probability distribution of the calibrated ^{14}C date; gray rhomb: the median; white square: the mean value.

In order to check which of the estimated parameters best estimate the true calendar age, the procedure is repeated for an assumed true calendar age that was changed every 20 yr in the range 0–14,000 cal BP. Calculations were carried out using the IntCal04 calibration curve (Reimer et al. 2004) and the revised and updated calibration module of the Gliwice Radiocarbon Laboratory calibration program GdCALIB (Aitchison et al. 1989; Pazdur and Michczyńska 1989; Michczyńska et al. 1990).

RESULTS AND DISCUSSION

Mode, Median, and Mean Value

The parameters that were tested initially were the mode, median, and mean of the probability distribution because these parameters are often used for estimating the calendar age. The mode is a value corresponding to the maximum of the density function, and the median is the value that divides the graph of the density function into 2 parts with the same probability ($= 0.5$). The mean (or expected)

value is calculated as a weighted average of all possible age values given by the density function, weighted by the probability that corresponds to that age. These 3 parameters have the same value for the unimodal Gaussian distribution, but in other cases their values become different (see Figure 1).

The results of the simulation are presented in Figure 2. Black triangles denote the difference between the true calendar age and mode of the probability distribution; gray rhombuses denote the difference between the true calendar age and median; and white squares indicate the difference between the true calendar age and the mean value. Figure 2 shows the results obtained for an assumed uncertainty value (σ) equal to 25 yr (Figure 2A), 50 yr (Figure 2B), and 100 yr (Figure 2C). It is clearly visible that for all tested parameters, large differences occur between the true calendar age and the value of the parameter, i.e. none of the tested parameters is a very good estimate of the true calendar age. We may also note that the values of calendar age estimated using the median and mean value almost always differ from the true calendar age, while the values of the calendar age estimated using the mode very often agree with the true calendar age. However, just for the mode we can sometimes observe the largest differences between the true and estimated calendar age.

Figure 3 shows the frequency of appearance of particular differences between the true and estimated calendar age of the sample in the range 0–14,000 cal BP for an assumed ^{14}C uncertainty equal to 25 yr (Figure 3A), 50 yr (Figure 3B), and 100 yr (Figure 3C). The black line represents the mode, dark gray line is the median, and light gray line is the mean value. We may observe that the differences for the median and mean value depend on the uncertainty of the ^{14}C date; they become smaller when the uncertainty decreases and larger when the uncertainty increases. For example, the differences between the true calendar age and the mean value are less than 35 yr for 66.2%, 54.9%, and 48.3% of the cases if the uncertainty is equal to 25, 50, and 100 yr, respectively. However, the differences for the mode are the same regardless of the uncertainty of the ^{14}C date. This shows that the reason for the differences is the shape of the calibration curve and the non-Normality of the resulting calibrated age density.

Figure 3 confirms that good concordance between the true and estimated calendar age appears most often when we estimate the calendar age using the mode of the probability distribution. The difference between the true calendar age and the mode of the probability distribution is equal to zero for 47.5%, and less than 35 yr for 73% of the cases.

Figure 4 answers the question of where the largest differences occur and shows how these differences may affect the chronology or time-depth model. We notice that the largest differences occur for those calendar years where the calibration curve has large wiggles and rather flat areas with steep parts at the ends. The probability distributions of calibrated ^{14}C dates are in these cases multimodal or flat and significantly differ from the Gaussian distribution. An influence of these differences on inference concerning chronology may be shown by the example of 2 samples with calendar ages equal to 2700 and 2400 cal BP (gray circles on the x axis in Figure 4). Depending on which estimating parameter we choose, for a sample with the true calendar age equal to 2700 cal BP we obtain the result shifted about 110–140 yr towards smaller age values, while for a sample with true calendar age equal to 2400 cal BP the results are shifted about 50–100 yr towards greater age values (shifts are marked in Figure 4 with arrows). Therefore, the estimated calendar age for both samples is almost the same (see Figure 4), and if we use these results for constructing a chronology or age-depth model we may draw wrong conclusions. The explanation for this effect is the flat area of the calibration curve occurring between 2400 and 2700 cal BP (Hallstadt period). It results in probability distributions for ^{14}C dates related to this period that are very wide, rather flat, and cover almost the whole period 2400–2700 cal BP, whereas the parameters that were applied here place themselves in the middle part of the distribution, even if the true calendar age is located at the end of the distribution.

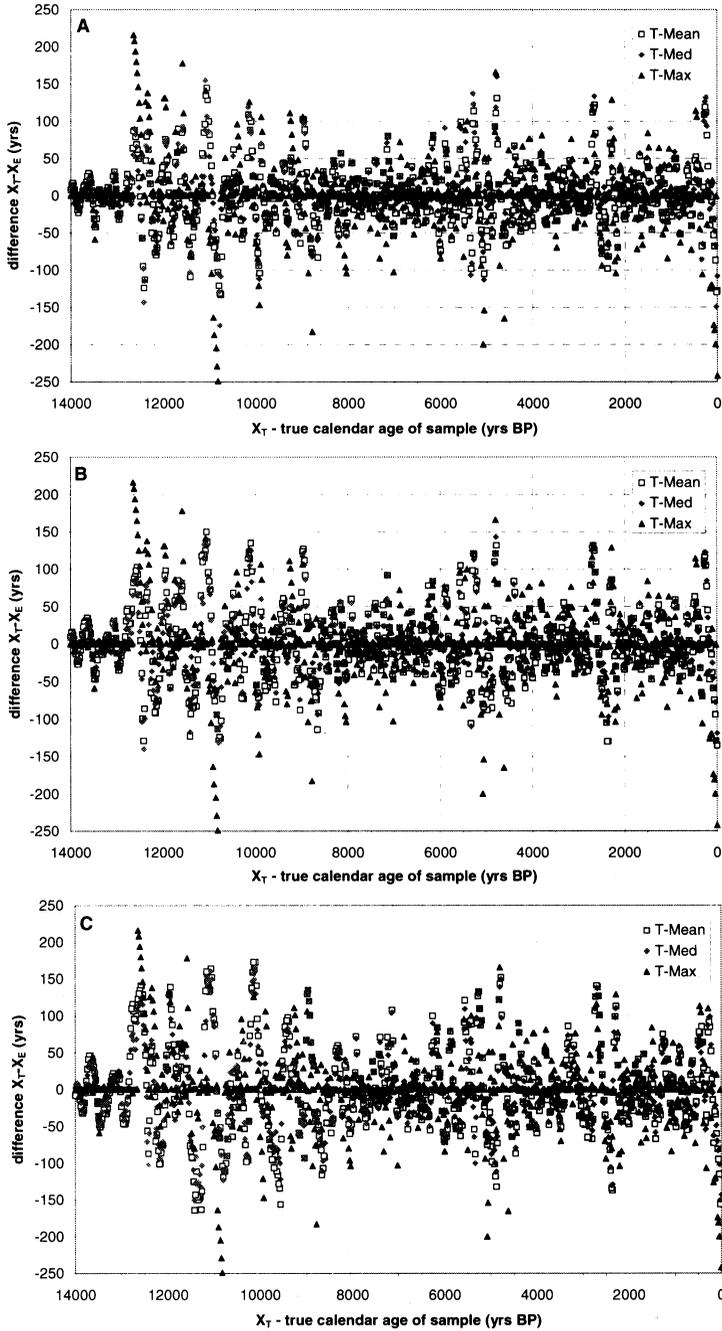


Figure 2 The differences between the true (X_T) and estimated (X_E) calendar age of a sample for the true calendar age of a sample changing in the range 0–14,000 cal BP. Results are obtained for the following values of the assumed uncertainty of a ^{14}C date: A: $\sigma = 25$ yr, B: $\sigma = 50$ yr, C: $\sigma = 100$ yr. Black triangles denote the difference between the true calendar age and the mode of the probability distribution; gray rhombuses are the difference between the true calendar age and the median; and white squares are the difference between the true calendar age and the mean value.

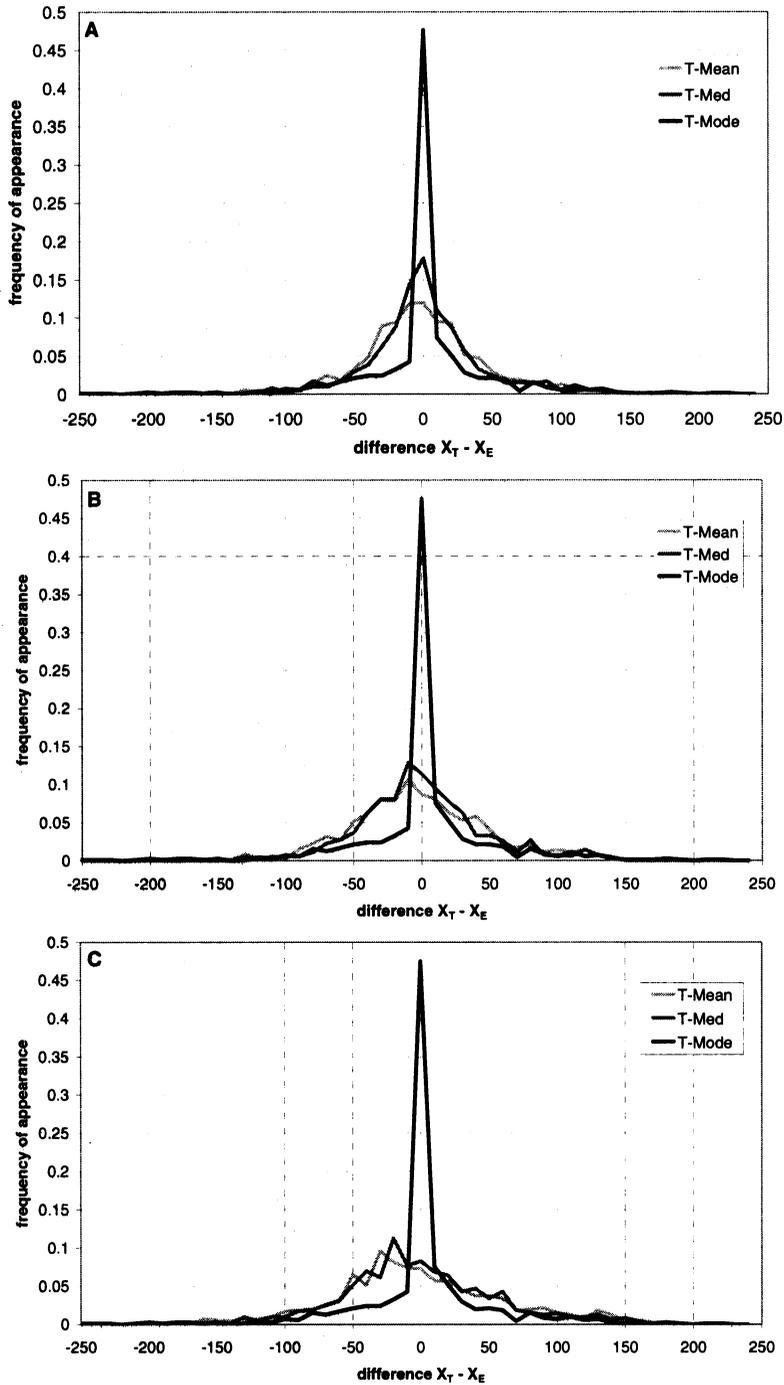


Figure 3 The frequency of appearance of particular differences between the true and estimated calendar age of a sample in the range 0–14,000 cal BP. The black line represents the mode; dark gray line: median; light gray line: mean value. The assumed uncertainty of ^{14}C date σ is equal to 25 yr (A), 50 yr (B), and 100 yr (C).

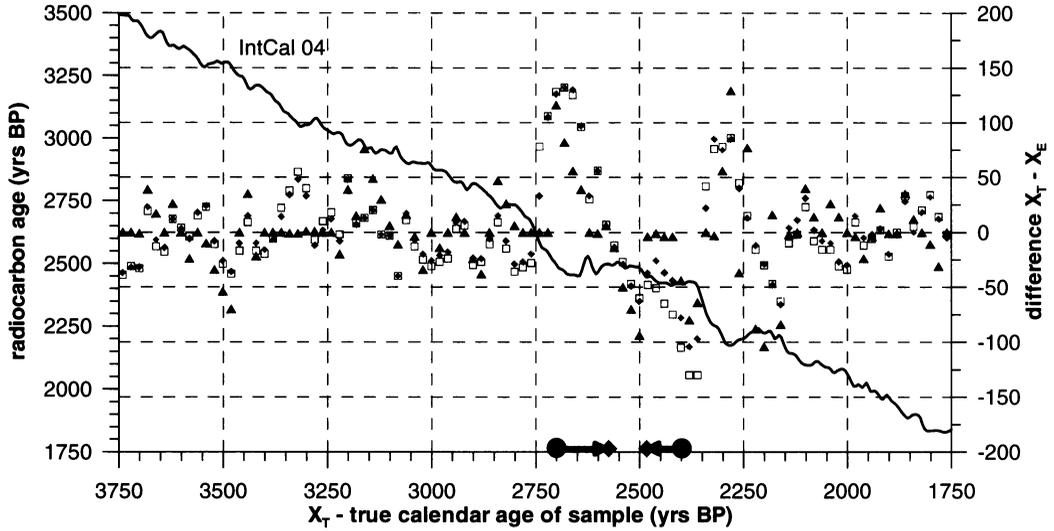


Figure 4 The differences between the true (X_T) and estimated (X_E) calendar age of a sample for the true calendar age of sample changing in the range 1750–3750 cal BP and the assumed uncertainty of ^{14}C date $\sigma = 50$ yr. The black line shows the calibration curve IntCal04 (Reimer et al. 2004). Black triangles denote the difference between the true calendar age and the mode of the probability density function; gray rhombuses are the difference between the true calendar age and the median; and white squares are the difference between the true calendar age and the mean value. The meaning of the symbols on the x axis is explained in the text.

Other Point Estimates

There are also other point estimates that are used often and should be discussed (see Figure 5):

- The local mode (maximum), i.e. age value, which corresponds to the maximum of the probability distribution included in the 68.2% or 95.4% confidence intervals (marked in Figure 5 with triangles);
- The central point (middle) of the confidence interval (marked in Figure 5 with asterisks).

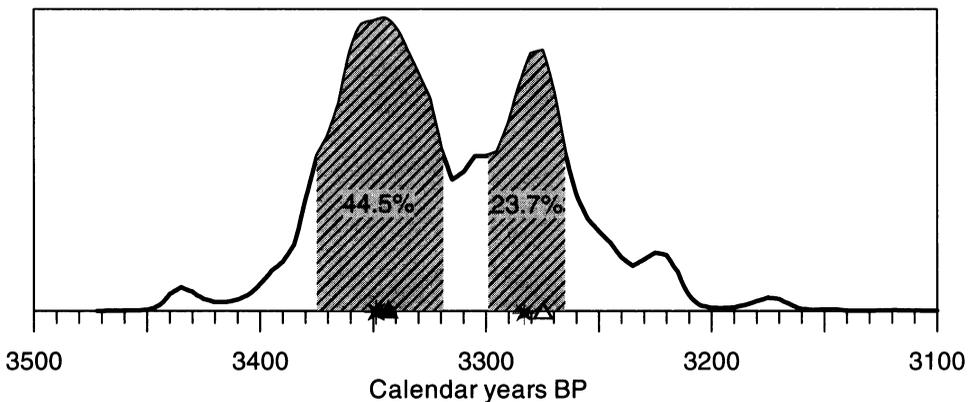


Figure 5 An example of 2 parts of the 68.2% confidence interval. Symbols on the x axis show the following: gray triangle: the mode of the first part of the interval; white triangle: the mode of the second part of the interval; black asterisk: the central point of the first part of the interval; and gray asterisk: the central point of the second part of the interval.

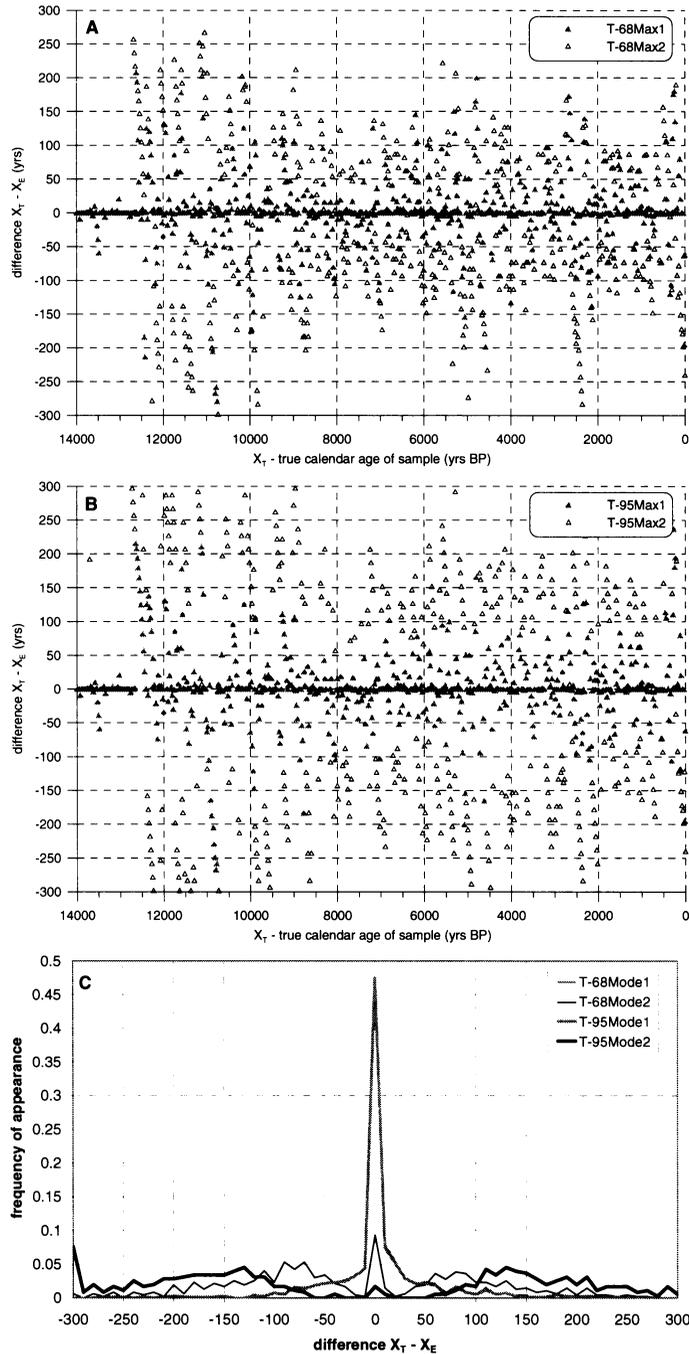


Figure 6 The differences between the true calendar age and the calendar age estimated using the mode of the first (gray triangles) and the second part (white triangles) of the 68.2% (A) and 95.4% (B) confidence intervals, assumed uncertainty of ^{14}C date $\sigma = 50$ yr. C) The frequency of appearance of particular differences. The thin black line represents the mode of the second part of the 68.2% interval; thick gray line: the mode of the first part of the 95.4% interval; thick black line: the mode of the second part of the 95.4% interval. The thin gray line, which should present the graph for the mode of the first part of the 68.2% interval, is not visible because this graph is identical with the graph for the mode of the first part of the 95.4% interval.

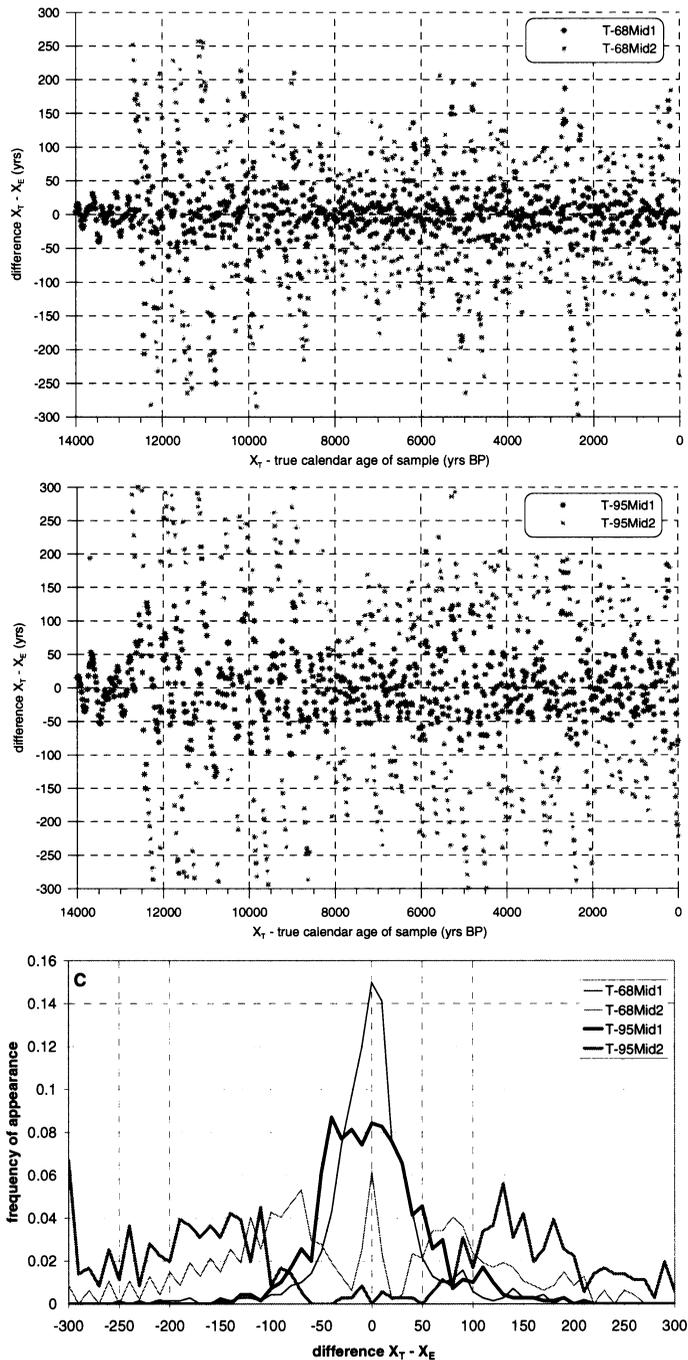


Figure 7 The differences between the true calendar age and the calendar age estimated using the central point of the first (black asterisks) and the second part (gray asterisks) of the 68.2% (A) and 95.4% (B) confidence intervals, assumed uncertainty of ^{14}C date $\sigma = 50$ yr. C) The frequency of appearance of particular differences. The thin black line represents the central point of the first part of the 68.2% interval; thin gray line: central point of the second part of the 68.2% interval; thick black line: central point of the first part of the 95.4% interval; thick gray line: central point of the second part of the 95.4% interval.

Because confidence intervals for calibrated ^{14}C dates are often divided into several parts, the above-mentioned point estimates were then calculated separately for each part. During the calculation, only 2 of the most important parts were taken into account: the most probable part of the confidence interval (which we term “the first part”) and the second most probable part (“the second part”).

Analysis was done for 68.2% and 95.4% confidence intervals and the assumed ^{14}C date uncertainty of 50 yr. The results are presented in Figures 6 and 7. We can see that the differences for the mode of the first part of the interval (Figure 6) are almost the same as the differences for the main mode (see Figure 3), because these modes are usually equivalent. The mode of the second part of the 68.2% confidence interval estimates the true calendar age of the sample for usually incorrectly, but sometimes it gives a correct value for the true calendar age of the sample (see Figures 6A and 6C). Figures 6B and 6C show, however, that the mode of the second part of a 95.4% confidence interval estimates the true calendar age incorrectly. The results presented in Figure 7 show that the central point of the first part of the 68.2% and 95.4% confidence intervals is a poor estimate of the true calendar age (due to the rather large differences between the true and estimated calendar age of the sample), and the central point of the second part is completely wrong as an estimate of the true calendar age.

CONCLUSIONS

The results of the simulation unambiguously show that a very good method for point estimation of the true calendar age of a sample does not exist. For all the tested parameters, one may observe differences between the true calendar age of a sample and the value of the calendar age estimated using the parameter. Therefore, the best method for presenting the calibrated ^{14}C dates is to use the whole probability density function.

If it is really essential to use a point estimate for the calendar age of the sample, then the mode (the value of the calendar age that corresponds to the maximum of the probability distribution of a calibrated ^{14}C date) may be accepted as a point estimate, but we should remember that important differences between the mode and the true calendar age of the sample appear for some periods, which are characterized by a specific shape of calibration curve (see Figure 4). The use of other parameters for the point estimation of calendar age may lead to wrong conclusions.

ACKNOWLEDGMENTS

This work was supported by statutory funds of the Silesian University of Technology BK-2006.

REFERENCES

- Aitchison TC, Leese M, Michczyńska DJ, Mook WG, Otle RL, Ottaway BS, Pazdur MF, van der Plicht J, Reimer PJ, Robinson SW, Scott EM, Stuiver M, Weninger B. 1989. A comparison of methods used for the calibration of radiocarbon dates. *Radiocarbon* 31(3):846–64.
- Michczyńska DJ, Pazdur MF, Walanus A. 1990. Bayesian approach to probabilistic calibration of radiocarbon ages. In: Mook WG, Waterbolk HT, editors. Proceedings of the 2nd International Symposium ^{14}C and Archaeology, Strasburg. *PACT* 29: 69–79.
- Pazdur MF, Michczyńska DJ. 1989. Improvement of the procedure for probabilistic calibration of radiocarbon dates. *Radiocarbon* 31(3):824–32.
- Reimer PJ, Baillie MGL, Bard E, Bayliss A, Beck JW, Bertrand CJH, Blackwell PG, Buck CE, Burr GS, Cutler KB, Damon PE, Edwards RL, Fairbanks RG, Friedrich M, Guilderson TP, Hogg AG, Hughen KA, Kromer B, McCormac G, Manning S, Bronk Ramsey C, Reimer RW, Remmele S, Southon JR, Stuiver M, Talamo S, Taylor FW, van der Plicht J, Weyhenmeyer CE. 2004. IntCal04 terrestrial radiocarbon age calibration, 0–26 cal kyr BP. *Radiocarbon* 46(3):1029–58.
- Stuiver M, Reimer P. 1989. Histograms obtained from computerized radiocarbon age calibration. *Radiocarbon* 31(3):817–32.
- Telford RJ, Heegaard E, Birks HJB. 2004. The intercept is a poor estimate of a calibrated radiocarbon age. *The Holocene* 14(2):296–8.