HOW SMOOTH SHOULD CURVES BE FOR CALIBRATING RADIOCARBON AGES?

TORBJÖRN E. TÖRNQVIST and MARC F. P. BIERKENS

Department of Physical Geography, Utrecht University, P.O. Box 80115, NL-3508 TC Utrecht
The Netherlands

ABSTRACT. We show that smoothed versions of the high-resolution calibration curve should be used when ^{14}C ages are calibrated with large (> ~30 ^{14}C yr) measurement errors (represented by standard deviation σ_m) or are mixtures of elements of variable age (natural sample error with standard deviation σ_n). The degree of smoothing should agree with the standard deviation of total sample error, σ_t, the square root of the quadratic sum of σ_m and σ_n. However, in most cases, σ_t is not well known, especially due to difficulties in quantifying σ_n. We present an inverse method that gives a measure of mean σ_t for different materials that are widely used in (conventional) ^{14}C dating. Calculations with large (>100) data sets of wood, charcoal, ombrotrophic peat and minerotrophic peat/gyttja samples indicate that σ_t of such materials is generally much larger than previously assumed, mainly because of large values of σ_n. This means that particularly in organic deposits, strongly smoothed calibration curves should be used where medium-term ^{14}C variations (wiggles) are completely straightened. This has especially major consequences for calibrating ^{14}C histograms for natural ^{14}C variations. We conclude that ^{14}C histograms consisting of samples of organic deposits do not require correction for medium-term ^{14}C variations and that uncalibrated ^{14}C histograms need not be as suspect as is usually believed.

INTRODUCTION

Calibration of ^{14}C ages has rapidly evolved into a standard procedure for ^{14}C users, facilitated by the development of numerous calibration programs (Aitchison *et al.* 1989). Calibration is usually performed with a high-resolution standard calibration curve, but it is questionable whether the use of such a detailed calibration curve is always justified. The calibration data of Stuiver and Pearson (1986), Pearson and Stuiver (1986) and Pearson *et al.* (1986), which have been the most widely used, are based on high-precision ^{14}C measurements of tree rings with very small standard deviations. Nearly all regular ^{14}C ages have significantly larger standard deviations, suggesting that, in such cases, the use of a smoothed version of the calibration curve may be more appropriate. A related problem in calibrating ^{14}C ages has been discussed by Mook, de Jong and Geertsema (1979), Mook (1983), Stuiver and Pearson (1986), Pearson and Stuiver (1986) and Michczyńska, Pazdur and Walanus (1990). This refers to the time range into which the sample fits (sample "time width" according to Mook 1983), which is also important for choosing a calibration curve. Because the high-resolution calibration curve is based on samples with small time widths (20 cal yr), it shows subtle ^{14}C variations that are probably not detectable in samples (*e.g.*, sediment samples) of significantly larger time width. Thus, the original calibration curve should be smoothed to suit the type of sample used before the calibration is carried out (Mook 1983).

We consider ^{14}C samples as populations of elements with different $^{14}C/^{12}C$ ratios (Appendix 1) containing two error sources: 1) natural sample error, ε_n, which is a measure of the degree of mixing of various elements of different age in the sample and is represented by standard deviation σ_n; and 2) measurement error, ε_m, which is represented by the standard deviation σ_m, as determined by the laboratory. The former is familiar to the concept of time width as introduced by Mook (1983). Time width refers to the number of calendar years represented by a sample. Only at times (notably with wood) is it possible to establish this accurately. A direct relation exists between time width and σ_n, as discussed in Appendix 2.

Most samples, such as organic deposits and charcoal, have unknown σ_n, which consists partly of the time width of the sample, and partly of various types of contamination. Note that our statistical assumption is that contamination will rarely lead to bimodal probability distributions. We expect

unimodal and more or less symmetrical probability distributions which can be described appropriately by the standard deviation. It is almost impossible to estimate contamination quantitatively for individual samples, but it may be significantly larger than the error resulting from sample time width alone. The standard deviation of total sample error, σ_t, must be established quantitatively to select the appropriate calibration curve for one ^{14}C sample or a set of ^{14}C samples. If σ_n is independent of σ_m (see Appendix 1), σ_t is obtained by (*cf.* Mook and Streurman 1983)

$$\sigma_t = \sqrt{(\sigma_n^{\,2} + \sigma_m^{\,2})} \quad [^{14}C \text{ yr}] . \tag{1}$$

Until now, it has been difficult to determine σ_t because of problems in determining σ_n. Thus, a suitable version of the calibration curve has usually been chosen by simply estimating mean time width of a sample with no quantitative analysis (*e.g.*, Mook 1983; Stolk, Hogervorst and Berendsen 1989). We present here an inverse method that estimates σ_t (and thus, σ_n, because σ_m is known) based on large (>100) data sets of various types of samples widely used in (conventional) ^{14}C dating. As will be shown, such an analysis has significant consequences for ^{14}C age calibration. Our results are particularly important for calibrating ^{14}C histograms for ^{14}C variations.

METHODS

"Real" and "Theoretical" ^{14}C Histograms

We selected published ^{14}C ages of different types of sample material and constructed ^{14}C histograms for each of these types (Fig. 1), using the CALHIS program (Stolk *et al.* 1994). ^{14}C histograms are obtained by superimposing individual ^{14}C ages, each represented by a Gaussian distribution. This resulted in a data set for each type of material consisting of intensity I (the height of the ^{14}C histogram) for every 10th ^{14}C yr. These four data sets are referred to as "real" ^{14}C histograms. Second, we calculated a parameter that represents the degree of clustering of ^{14}C ages due to natural ^{14}C variations that can be expected for samples equally distributed over the calendar time scale. This parameter, the C-value, is calculated for the same segments of 10 ^{14}C yr as in the "real" ^{14}C histograms, with the equation (see also Fig. 2)

$$C = \Delta \text{ cal yr} / \Delta\,^{14}C \text{ yr} \qquad \text{where } \Delta\,^{14}C \text{ yr} = 10. \tag{2}$$

We used 25 calibration curves, each with different degrees of smoothing, for this analysis (Fig. 3A). We smoothed the curves using the SMOOTH program, which is part of the Groningen calibration package (van der Plicht and Mook 1989; van der Plicht 1993). This program calculates moving averages over the calendar time scale using a Gauss-function for weighting. Our most detailed calibration curve (calculated with a Gauss-function with smoothing parameter $\sigma_s = 8$) contains all the details of the high-resolution calibration curve, whereas a curve smoothed with $\sigma_s = 200$ only registers the long-term ^{14}C variations (wiggles have been completely straightened; Fig. 3A).

A plot of C-values *vs.* ^{14}C age (Fig. 3B) can be regarded as a "theoretical" ^{14}C histogram that is expected when the distribution of ^{14}C samples over the calendar time scale is homogeneous. Peaks and troughs in such a histogram result entirely from ^{14}C variations. Because C-values were calculated for each of the 25 calibration curves, we obtained 25 "theoretical" ^{14}C histograms, each corresponding to a calibration curve with a different σ_s value. Our method performs a quantitative comparison of the ^{14}C histograms constructed for different data sets ("real" ^{14}C histograms) and this sequence of "theoretical" ^{14}C histograms. The idea behind this comparison is that the "theoretical" ^{14}C histogram that shows the best correspondence with a "real" ^{14}C histogram represents the cali-
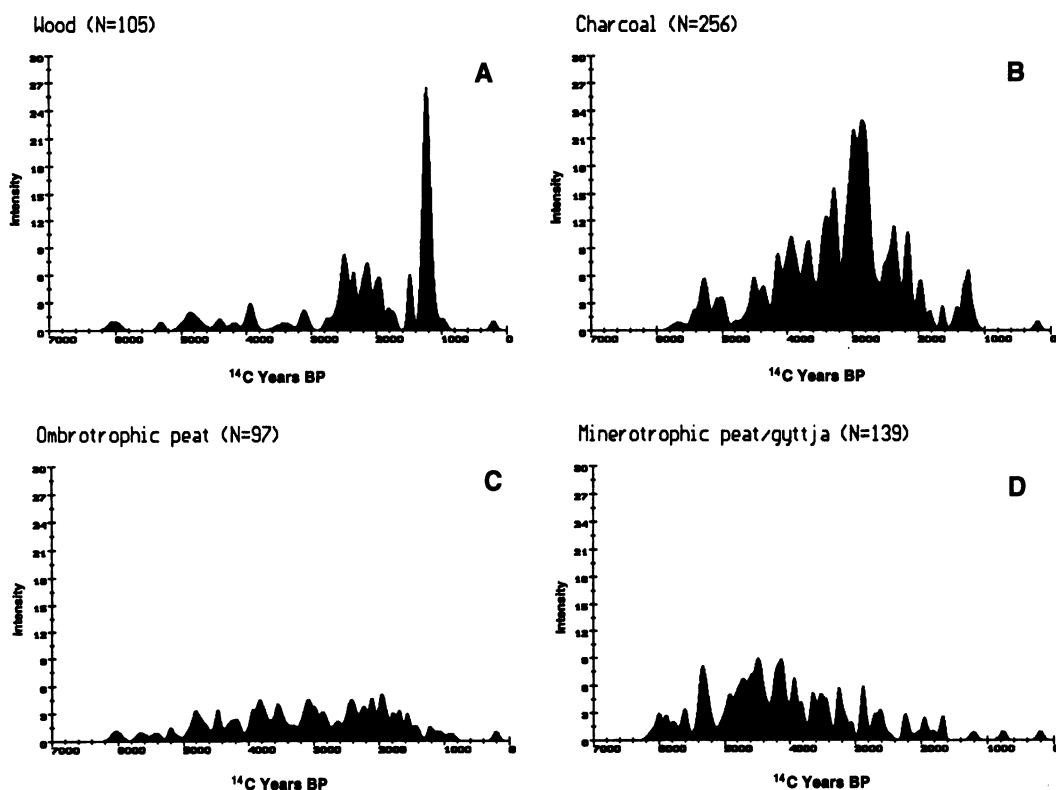
Fig. 1. [14]C histograms of data sets of (A) wood, (B) charcoal, (C) ombrotrophic peat and (D) minerotrophic peat/gyttja, constructed with the CALHIS program (Stolk *et al.* 1994). Standard peak at 200 yr BP represents one ^{14}C age with I = 1 and σ_m = 45. The raw data were obtained from various sources; wood and charcoal: Lanting and Mook (1977); ombrotrophic peat: Janssen and ten Hove (1971), Casparie (1972), van Geel (1972, 1978), Middeldorp (1982, 1986), Dupont and Brenninkmeijer (1984), Kuhry (1985), Witte and van Geel (1985), Dupont (1986); minerotrophic peat/gyttja: Berendsen (1982), van der Woude (1983), van Dijk, Berendsen and Roeleveld (1991), Törnqvist and van Dijk (1993).
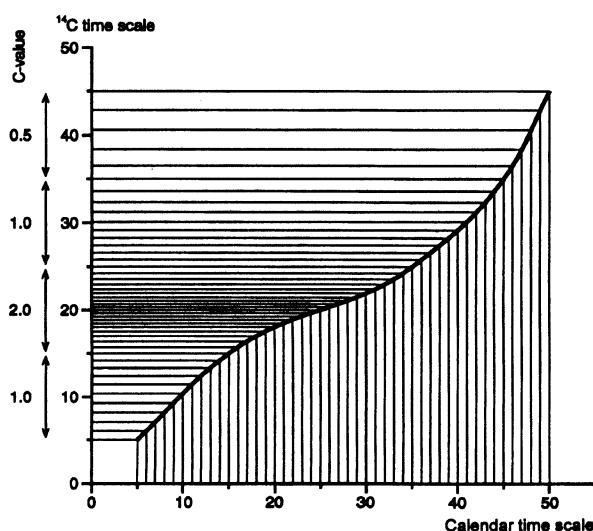

Fig. 2. Determination of the C-value. ——— = a theoretical calibration curve; ——— = fictitious ^{14}C ages for each calendar year unequally distributed over the ^{14}C time scale due to ^{14}C variations. The C-value is determined for every 10th ^{14}C yr.

bration curve that should be used for the type of sample from which the "real" [14]C histogram is constructed.

Determining Correspondence between "Real" and "Theoretical" [14]C Histograms

Because the range of C-values is different for each calibration curve (Fig. 3B), we subdivided the C-values into eight classes of equal size (C-class 1 being the smallest 12.5% of the C-values, C-class 8 the largest; Figs. 3C–E). First we determined the intensity of each C-class in the "real" [14]C histogram, which revealed the observed intensity for each C-class (I_i^{obs}; see Appendix 3). If there is no
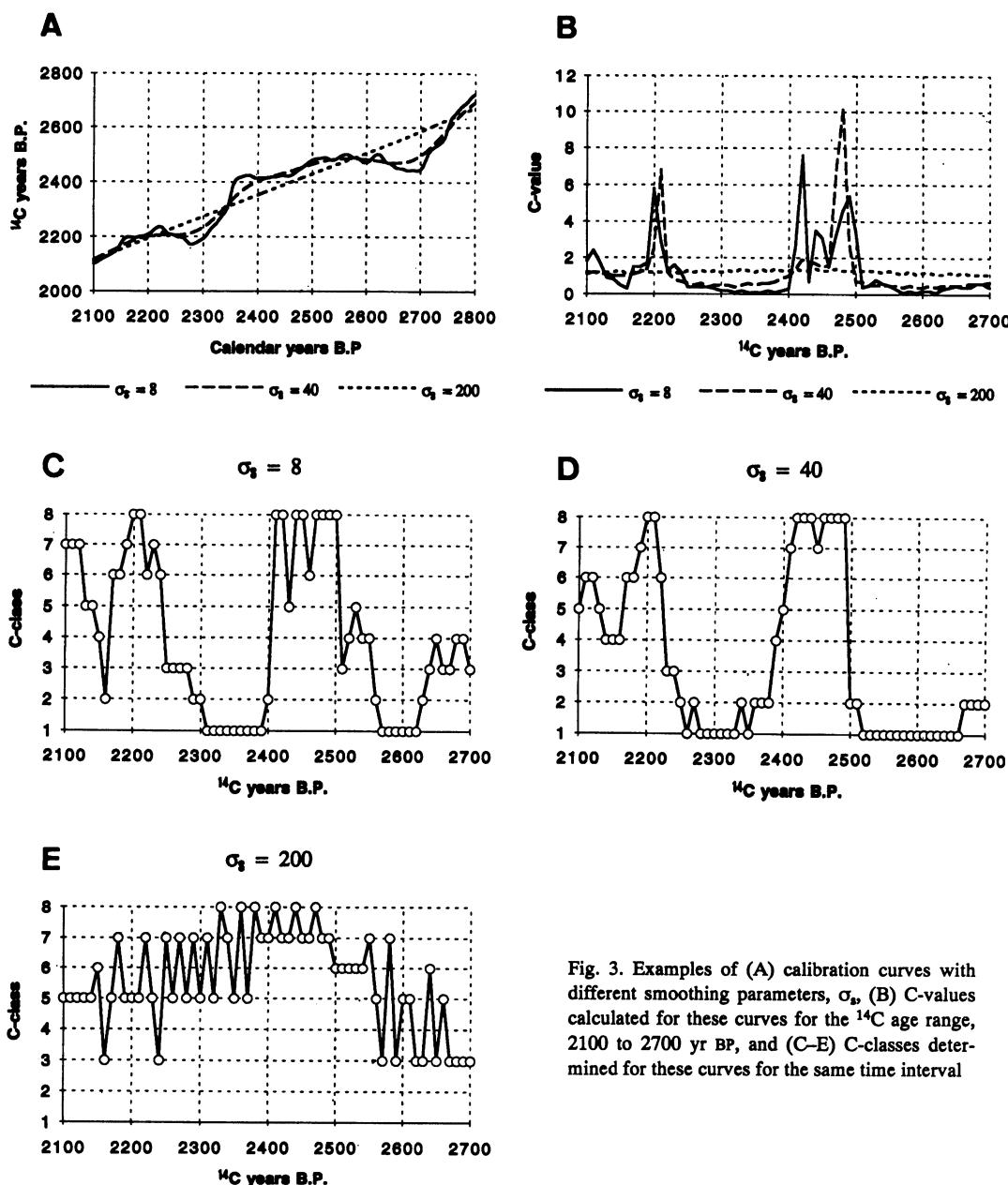

Fig. 3. Examples of (A) calibration curves with different smoothing parameters, σ_s, (B) C-values calculated for these curves for the ^{14}C age range, 2100 to 2700 yr BP, and (C–E) C-classes determined for these curves for the same time interval

relation between the "real" ^{14}C histogram and the "theoretical" ^{14}C histogram, we expect an equal intensity for each C-class. This is the intensity that can be expected in each C-class in the case of random distribution of ^{14}C ages (I_i^{exp}), without any clustering due to ^{14}C variations. In principle, I_i^{exp} can be calculated easily for each C-class by dividing the total intensity by the number of C-classes. However, the unequal large-scale distribution of ^{14}C ages over the ^{14}C time scale in most of our data sets (Fig. 1), combined with the irregular distribution of the C-class numbers over the ^{14}C time scale, require a weighted calculation, as discussed in Appendix 3.

If ^{14}C ages cluster in certain time intervals due to ^{14}C variations, indicating a relation between the "real" and the "theoretical" ^{14}C histogram, the ratio I_i^{obs}/I_i^{exp} will be <1 for the lower C-classes and >1 for the higher C-classes. If the I_i^{obs}/I_i^{exp} ratio is plotted *vs.* C-class number, the slope, a, of a regression line through the data points is a measure of the degree of correspondence between the "real" and the "theoretical" ^{14}C histograms for the calibration curve used (Fig. 4).
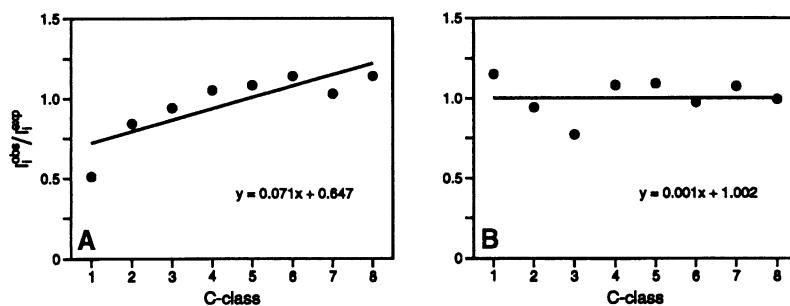

Fig. 4. Scatterplot of the I_i^{obs}/I_i^{exp} ratio *vs.* C-class number and a linear regression line through the data points. Examples are given of the result for charcoal with (A) a calibration curve with σ_s = 96 and (B) a calibration curve with σ_s = 200. The former shows a significant relation between I_i^{obs}/I_i^{exp} ratio and C-class, the latter not.

Estimating σ_t

When samples with large values of σ_t are calibrated with a calibration curve that is too detailed, the resulting probability distributions will contain artifacts. On the other hand, a curve that is too smooth will not correct sufficiently for ^{14}C variations, which actually may be detectable in ^{14}C samples. Thus, the "theoretical" ^{14}C histogram that corresponds the best to the "real" ^{14}C histogram represents the most suitable calibration curve for that specific type of sample; it will correct for all relevant variation without producing artifacts. The σ_s value with the maximum slope of the I_i^{obs}/I_i^{exp} ratio is a measure for σ_t and *vice versa*. The relation between σ_t and σ_s is complex because σ_s is expressed in calendar years and σ_t in ^{14}C years. However, a conversion method is presented in Appendix 4, which can be used for estimating σ_t from σ_s for larger data sets. By plotting the slope, a, of the I_i^{obs}/I_i^{exp} ratio regression line *vs.* the smoothing parameter, σ_s, of the calibration curves, we estimate σ_t by means of an inverse method.

RESULTS

Experiments

We applied the procedure described above first to test the method. We used the set of ^{14}C ages that formed the basis for the standard calibration curve (Stuiver and Pearson 1986; Pearson and Stuiver 1986; Pearson *et al.* 1986) and calculated a σ_s-a plot for their high-precision wood samples. The

result (Fig. 5) is as expected: a steadily decreasing slope of the I_i^{obs}/I_i^{exp} ratio with increased smoothing of the calibration curve. The maximum value is found at the most detailed calibration curve (σ_s = 8). Obviously, this is because mean σ_m is only 12.7 ^{14}C yr (Table 1) and σ_n is also very small, since the time width covered by the samples is only 20 cal yr.

Stuiver/Pearson data (N=362)

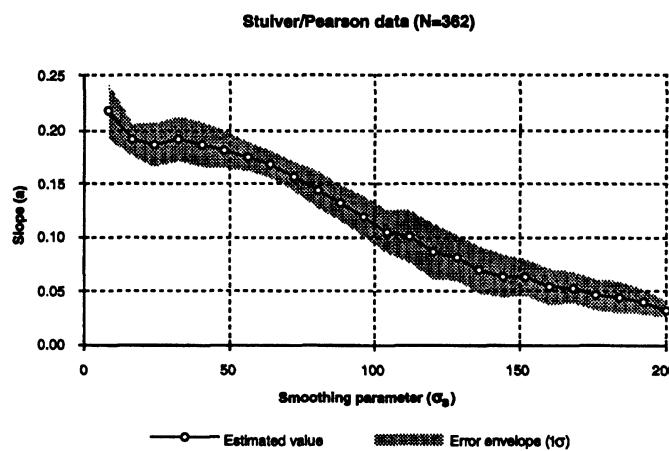

Fig. 5. σ_s-a plot of the Stuiver and Pearson (1986), Pearson and Stuiver (1986) and Pearson *et al.* (1986) data set

TABLE 1. Mean Measurement Error, σ_m, of Different Data Sets Used in This Study

Data set	N	Mean σ_m (^{14}C yr)
Stuiver/Pearson data	362	12.7 ± 3.7
Wood	105	46.2 ± 21.1
Charcoal	256	46.8 ± 16.8
Ombrotrophic peat	97	56.4 ± 19.1
Minerotrophic peat/gyttja	139	59.3 ± 19.4
Charcoal (high σ_m)	100	62.9 ± 16.1
Charcoal (low σ_m)	100	33.8 ± 2.5

Before analyzing and comparing data sets of different materials, we also investigated the number of ^{14}C ages required to arrive at reproducible results. We used charcoal samples (N = 256) and selected 3 random sets of 50 and 3 of 100 ^{14}C ages. In neither random set did mean σ_m differ significantly from mean σ_m of the total data set. A comparison of σ_s-a plots of the total charcoal data set with the 50 (Fig. 6A) and 100 ages (Fig. 6B) showed that data sets of 50 were too small, indicated by the large differences among the curves of Figure 6A. Not only did their absolute values differ, but their shapes especially differed (maximum values of a are reached at different values of σ_s). Figure 6B shows that sets of 100 yield much better results; the shape of the three curves strongly resembles the "total charcoal" curve. This suggests that data sets of >100 are needed to obtain useful σ_s-a plots.

We also analyzed the effect of different σ_m values. Again, we used our charcoal data set, and split it into two subsets of 100 samples each; one with low standard deviations and one with high standard deviations (Table 1). A comparison of the two, together with the "total charcoal" curve (Fig. 7) confirms that σ_m strongly influences σ_t. Maximum values for charcoal samples with low σ_m are found

Effect of smaller data sets (N=50)
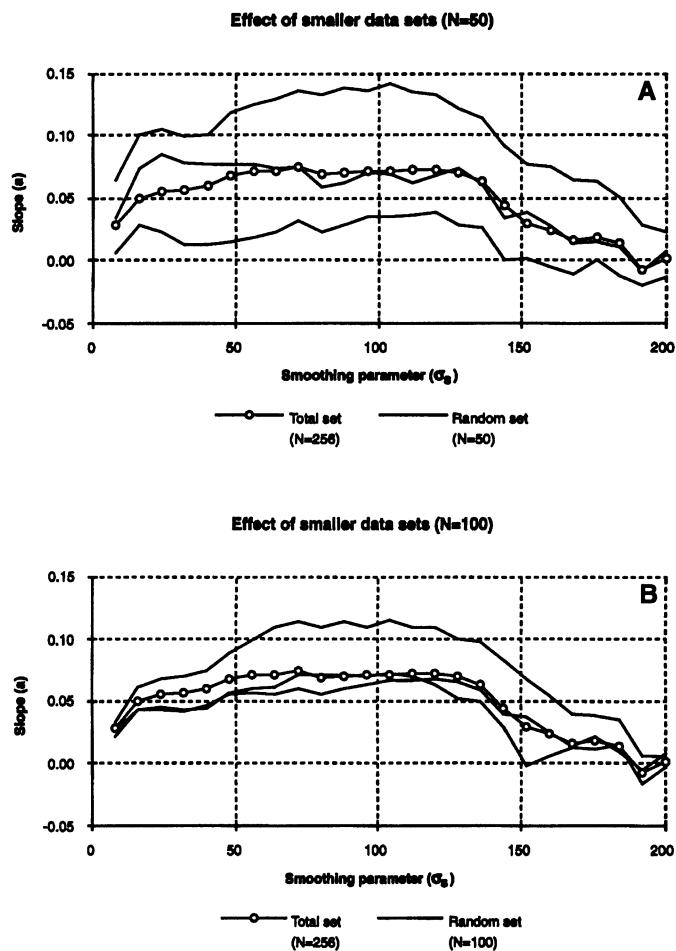

Effect of smaller data sets (N=100)

Fig. 6. σ_s-a plots of various data sets of charcoal. Comparison of data sets of (A) 50 samples each and (B) 100 samples each with the total charcoal data set

at smoothing parameters of *ca.* 60; for samples with high σ_m, maximum values are found at *ca.* 120. This confirms that smoothed curves are needed when samples with high standard deviations are calibrated.

Analysis of Different Materials

Data Selection

We critically selected ^{14}C ages of four types of material widely used in ^{14}C dating: wood, charcoal, ombrotrophic peat and minerotrophic peat/gyttja. We used ^{14}C results from only one laboratory (Groningen), to rule out possible interlaboratory differences (International Study Group 1982; Scott *et al.* 1990). Further, we only selected samples with numbers higher than GrN-4046, because these were corrected for isotope fractionation (Lanting and Mook 1977). We restricted our sample selection to the time interval, 110–6100 BP, the period covered by the calibration data of Stuiver and Pearson (1986), Pearson and Stuiver (1986) and Pearson *et al.* (1986).
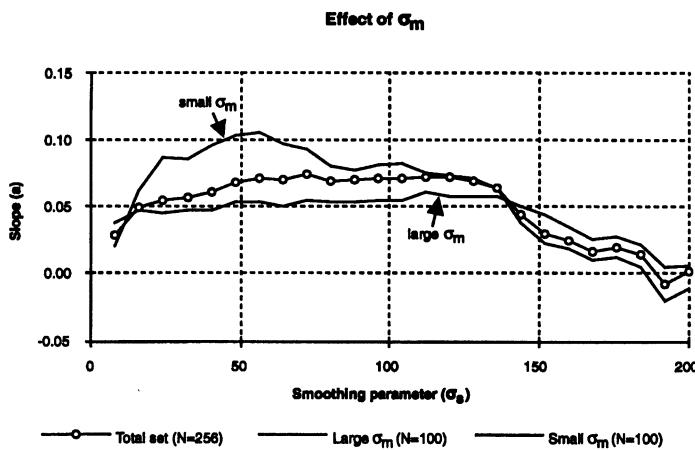
Fig. 7. σ_s-a plots of various data sets of charcoal. Comparison of a data set with low σ_m and high σ_m with the total charcoal data set

Wood and charcoal samples were archaeological samples from The Netherlands. The ombrotrophic peat samples were obtained from raised bogs, mainly in the eastern Netherlands. The minerotrophic peat/gyttja samples, which usually have a significant admixture of clay, are all from the Rhine-Meuse delta (central Netherlands). To obtain a homogeneous data set, we used only minerotrophic peat/gyttja samples that had not been manually pretreated (*e.g.*, by systematic removal of roots). Figure 1 shows ^{14}C histograms of the 4 data sets.

Results and Interpretation

Figure 8 shows the results of all our analyses. The σ_s-a plots of wood (Fig. 8A) and charcoal (Fig. 8B) show significant maximum values of the slope of the I_i^{obs}/I_i^{exp} ratio. This is not the case in the σ_s-a plots of the organic deposits (Fig. 8C, D), where a horizontal line can be drawn through the error envelopes. The maximum values in Figures 8A and 8B do not occur as distinct peaks, as they do in the Stuiver/Pearson data of Figure 5, but as plateaus. This is not surprising; the variability of σ_m for these data sets is much higher than for the Stuiver/Pearson data (Table 1). The range of maximum values is especially wide in the charcoal curve (σ_s = 48–136, mean 92). Thus, we conclude that the variability of σ_n in the charcoal samples contributes much to this phenomenon. Mean σ_n can be calculated easily with the equation

$$\sigma_n = \sqrt{(\sigma_t^2 - \sigma_m^2)} \quad [^{14}C \text{ yr}]. \tag{3}$$

Determining σ_n from σ_s-a plots requires that σ_t (^{14}C yr) is estimated from σ_s (cal yr). Using the method discussed in Appendix 4, the correction factor for the conversion of σ_s into σ_t (1/Q) was calculated for the ranges 2800–1000 BP (wood) and 4600–1800 BP (charcoal), the time intervals that contained most of the ^{14}C ages of these materials (Fig. 1). Using Equation 3, mean σ_n for charcoal can be estimated at *ca.* 50, whereas for wood (range of maximum values at σ_s = 24–72, mean 48), the value is close to 0. Note that this calculation is only valid if the laboratory error multiplier (Stuiver and Pearson 1986; Pearson and Stuiver 1986) is close to 1.0. If the laboratory error multiplier is large, the contribution of σ_n to σ_t is overestimated.

The situation is rather different for the data sets of ombrotrophic peat and minerotrophic peat/gyttja. The σ_s-a plots (Fig. 8C, D) not only show much lower values of a, they are also more or less flat.
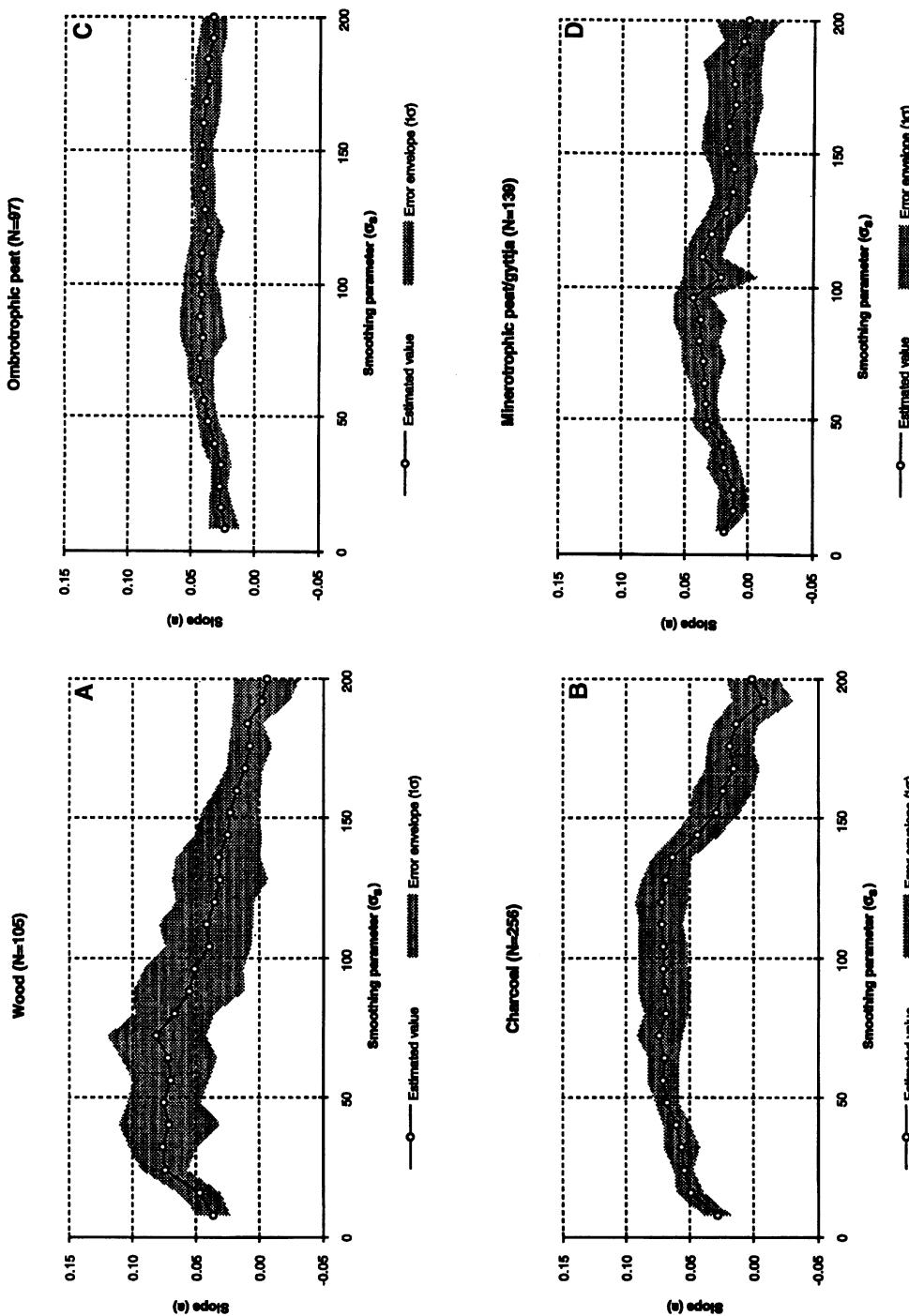
Fig. 8. σ_s-a plots for data sets of ^{14}C ages of (A) wood, (B) charcoal, (C) ombrotrophic peat and (D) minterotrophic peat/gytta. Error envelope represents the 1 σ confidence interval of the value of a.

Weak maxima seem to occur at values of σ_s of 60–100, but virtually all results are within each other's error limits, and thus not statistically significant. The conclusion is that medium-term ^{14}C variations cannot be detected in large data sets of bulk ^{14}C samples of these materials.

DISCUSSION AND RECOMMENDATIONS

General Implications for Calibration

Our experimental results show that σ_n (the standard deviation of natural sample error) and σ_m (the standard deviation of measurement error) strongly influence the degree at which medium-term ^{14}C variations are detectable in ^{14}C samples. As mentioned above, curves for calibrating ^{14}C ages need to agree with the standard deviation of total sample error (σ_t), in order to arrive at correct calibration results. The degree of smoothing of calibration curves depends on the type of material. Until now, appropriate calibration curves have been chosen on the basis of a "best guess" approach; for example, for sediment samples, 100-cal-yr moving averages of the high-resolution curve have been proposed (Mook 1983; Stuiver and Pearson 1986; Pearson and Stuiver 1986; Stolk, Hogervorst and Berendsen 1989). Our results show that a 100-cal-yr moving average smoothed version of the high-resolution calibration curve (comparable to our $\sigma_s = 40$ curve) will usually be too low an estimate. Especially organic deposits require a much smoother curve, which only reflects long-term ^{14}C variations.

The results obtained for organic deposits are not very surprising in view of recent findings that contamination of bulk minerotrophic peat/gyttja samples may produce age deviations on the order of several centuries (Törnqvist *et al.* 1992). The results for wood and charcoal also agree with expectations. The low σ_n determined for wood samples confirms the general opinion that wood yields excellent ^{14}C ages because of its well-defined and usually narrow time width, and because it is unlikely to be contaminated. More problems occur with charcoal samples, which are more likely to contain contaminants. Further, σ_n of charcoal is very variable, according to our results.

In σ_s-a plots with maximum values as plateaus, for example, in our charcoal data set (Fig. 8B), the choice of a suitable calibration curve is ambiguous. If one uses a calibration curve that is too smooth, probability distributions are not sufficiently corrected for the influence of medium-term ^{14}C variations. On the other hand, a curve that is too detailed will reveal irregular probability distributions consisting of narrow and high peaks, which are, in fact, artifacts. Both situations should be avoided. To eliminate ambiguity, we recommend using the mean value of σ_s in the plateau. We realize that this is a subjective choice, but the use of the mean value is most likely the best way to reduce potential error as much as possible. For our charcoal data set, $\sigma_s = 100$ is appropriate; for our wood samples, the value should be 50. For our peat samples, any calibration curve reflecting medium-term ^{14}C variations is too detailed; here, the completely smooth $\sigma_s = 200$ curve is suitable for avoiding artifacts. In general, for the proper calibration of ^{14}C ages, calibration programs need to have a smoothing option and the smoothing method needs to be described clearly.

When one calibrates results of other types of material where large data sets are not available, the method presented here cannot be used. If sample time width (expressed in cal yr) is accurately known or can be estimated reasonably well, it may be necessary to calculate σ_n (expressed in ^{14}C yr) from the time width value, to arrive at an acceptable estimate of σ_t. A calculation method is presented in Appendix 2.

The inverse method used here yields only average results for large data sets, which, as has been shown, usually are very variable. Thus, the first step should be an evaluation of all relevant evi-

dence available for individual samples, before considering the use of the results of this inevitably imprecise method.

Implications for Calibration of [14]C Histograms

When individual ^{14}C ages are calibrated, interest is usually focused on the calendar age range rather than the shape of the probability distribution. In such cases, the calibration result is usually not very sensitive to the degree of smoothing of the calibration curve. However, if the probability distribution of the calibrated ^{14}C age is important, the selection of a calibration curve becomes critical. The use of high-resolution calibration curves results in irregular calendar age probability distributions, whereas strongly smoothed calibration curves will result in probability distributions that remain generally Gaussian. This has important consequences for the calibration of ^{14}C histograms for ^{14}C variations. When calibrated ^{14}C histograms consist of summed probability distributions of individually calibrated ^{14}C ages (Stuiver and Reimer 1989; Stolk *et al.* 1994) the choice of the calibration curve is most crucial.

It has long been recognized that care should be taken with interpreting ^{14}C histograms that have not been corrected for ^{14}C variations (*e.g.*, Geyh 1971; de Jong and Mook 1981), resulting in a general suspicion of uncalibrated ^{14}C histograms. Peaks and troughs in uncalibrated ^{14}C histograms were thought to be the result of medium-term ^{14}C variations rather than the phenomena of interest. Our results show that calibration of ^{14}C histograms for medium-term ^{14}C variations may not be necessary, depending on the type of material. In the case of organic deposits, which have been widely used for constructing ^{14}C histograms, disturbing effects of medium-term ^{14}C variations are unlikely, and uncalibrated ^{14}C histograms are, in principle, very useful. On the other hand, it must be kept in mind that ^{14}C histograms based on bulk organic sediment samples are likely to register contaminating effects (*cf.* Törnqvist *et al.* 1992), and thus may have limited value.

Earlier attempts to correct ^{14}C histograms for ^{14}C variations (Geyh 1971, 1980) revealed synchroneity of peaks in the calibrated histograms with distortions of the ^{14}C time scale. Geyh (1971) assumed a causal relation between medium-term ^{14}C variations (caused by fluctuations in solar activity) and global sea-level changes, although he was later less certain about this relation (Geyh 1980). Our results indicate that peaks in Geyh's calibrated ^{14}C histograms may well be artifacts, resulting from the use of too-detailed calibration curves. Similarly, the large peaks produced by the KORHIS program for calibration of ^{14}C histograms (Stolk, Hogervorst and Berendsen 1989) are also artifacts, as shown by Stolk *et al.* (1994). Thus, we believe that studies on the relation between medium-term ^{14}C variations and geologic or paleoecologic phenomena should be carried out by wiggle matching (*cf.* van Geel and Mook 1989).

REFERENCES

Aitchison, T. C., Leese, M., Michczynska, D. J., Mook, W. G., Otlet, R. L., Ottaway, 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., Kra, R. S. and Srdoč, D., eds., Proceedings of the 13th International [14]C Conference. *Radiocarbon* 31(3): 846–862.

Berendsen, H. J. A. 1982 De genese van het landschap in het zuiden van de provincie Utrecht, een fysisch-geo-

grafische studie. *Utrechtse Geografische Studies* 25: 1–255.

Casparie, W. A. 1972 *Bog Development in Southeastern Drenthe (The Netherlands)*. The Hague, W. Junk: 271 p.

de Jong, A. F. M. and Mook, W. G. 1981 Natural C-14 variations and consequences for sea-level fluctuations and frequency analysis of periods of peat growth. *In* van Loon, A. J., ed., Quaternary Geology: A Farewell to A. J. Wiggers. *Geologie en Mijnbouw* 60: 331–336.

Dupont, L. M. 1986 Temperature and rainfall variation in the Holocene based on comparative palaeoecology and isotope geology of a hummock and a hollow (Bourtangerveen, The Netherlands). *Review of Palaeobotany and Palynology* 48: 71–159.

Dupont, L. M. and Brenninkmeijer, C. A. M. 1984 Palaeobotanic and isotopic analysis of late Subboreal and early Subatlantic peat from Engbertsdijksveen VII, The Netherlands. *Review of Palaeobotany and Palynology* 41: 241–271.

Geyh, M. A. 1971 Middle and young Holocene sea-level changes as global contemporary events. *Geologiska Föreningens i Stockholm Förhandlingar* 93: 679–692.

_____1980 Holocene sea-level history: Case study of the statistical evaluation of ^{14}C dates. *In* Stuiver, M. and Kra, R. S., eds., Proceedings of the 10th International ^{14}C Conference. *Radiocarbon* 22(3): 695–704.

International Study Group 1982 An inter-laboratory comparison of radiocarbon measurements in tree rings. *Nature* 298: 619–623.

Janssen, C. R. and ten Hove, H. A. 1971 Some Late Holocene pollen diagrams from the Peel raised bogs (southern Netherlands). *Review of Palaeobotany and Palynology* 11: 7–53.

Kuhry, P. 1985 Transgression of a raised bog across a coversand ridge originally covered with an oak-lime forest. Palaeoecological study of a Middle Holocene local vegetational succession in the Amtsven (northwest Germany). *Review of Palaeobotany and Palynology* 44: 303–353.

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. 1990 Bayesian approach to probabilistic calibration of radiocarbon ages. *In* Mook, W. G. and Waterbolk, H. T., eds., *Proceedings of the 2nd International Symposium ^{14}C and Archaeology*. Strasbourg, PACT 29: 69–79.

Middeldorp, A. A. 1982 Pollen concentration as a basis for indirect dating and quantifying net organic and fungal production in a peat bog ecosystem. *Review of Palaeobotany and Palynology* 37: 225–282.

_____1986 Functional palaeoecology of the Hahnenmoor raised bog ecosystem – A study of vegetation history, production and decomposition by means of pollen density dating. *Review of Palaeobotany and Palynology* 49: 1–73.

Mook, W. G. 1983 ^{14}C calibration curves depending on sample time-width. *In* Mook, W. G. and Waterbolk, H. T., eds., *Proceedings of the International Symposium ^{14}C and Archaeology*. Strasbourg, PACT 8: 517–525.

Mook, W. G., de Jong, A. F. M. and Geertsema, H. 1979 Archaeological implications of natural carbon-14 variations. *Palaeohistoria* 21: 9–18

Mook, W. G. and Streurman, H. J. 1983 Physical and chemical aspects of radiocarbon dating. *In* Mook, W. G. and Waterbolk, H. T., eds., *Proceedings of the International Symposium ^{14}C and Archaeology*. Strasbourg, PACT 8: 31–55.

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 1840 to 5210 BC. *In* Stuiver, M. and Kra, R. S., eds., Proceedings of the 12th International ^{14}C Conference. *Radiocarbon* 28(2B): 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., Proceedings of the 12th International ^{14}C Conference. *Radiocarbon* 28(2B): 839–862.

Scott, E. M., Aitchison, T. C., Harkness, D. D., Cook, G. T. and Baxter, M. S. 1990 An overview of all three stages of the International Radiocarbon Intercomparison. *In* Scott, E. M., Long, A. and Kra, R. S., eds., Proceedings of the International Workshop on Intercomparison of ^{14}C Laboratories. *Radiocarbon* 32(3): 309–319.

Stolk, A., Hogervorst, K. and Berendsen, H. 1989 Correcting ^{14}C histograms for the non-linearity of the radiocarbon time scale. *Radiocarbon* 31(2): 169–177.

Stolk, A., Törnqvist, T. E., Hekhuis, K. P. V., Berendsen, H. J. A. and van der Plicht, J. 1994 Calibration of ^{14}C histograms: A comparison of methods. *Radiocarbon* 36(1): 1–10.

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., Proceedings of the 12th International ^{14}C Conference. *Radiocarbon* 28(2B): 805–838.

Stuiver, M. and Reimer, P. 1989 Histograms obtained from computerized radiocarbon age calibration. *In* Long, A., Kra, R. S. and Srdoč, D., eds., Proceedings of the 13th International ^{14}C Conference. *Radiocarbon* 31(3): 817–823.

Törnqvist, T. E. and van Dijk, G. J. 1993 Optimizing sampling strategy for radiocarbon dating of Holocene fluvial systems in a vertically aggrading setting. *Boreas* 22: 129–145.

Törnqvist, T. E., de Jong, A. F. M., Oosterbaan, W. A. and van der Borg, K. 1992 Accurate dating of organic deposits by AMS ^{14}C measurement of macrofossils. *In* Long, A. and Kra, R. S., eds., Proceedings of the 14th International ^{14}C Conference. *Radiocarbon* 34(3): 566–577.

van der Plicht, J. 1993 The Groningen radiocarbon cali-
 bration program. *In* Stuiver, M., Long, A. and Kra, R.
 S., eds., Calibration 1993. *Radiocarbon* 35(1): 231–
 237.
van der Plicht, J. and Mook, W. G. 1989 Calibration of ra-
 diocarbon ages by computer. *In* Long, A., Kra, R. S.
 and Srdoč, D., eds., Proceedings of the 13th Interna-
 tional [14]C Conference. *Radiocarbon* 31(3): 805–816.
van der Woude, J. D. 1983 Holocene paleoenvironmental
 evolution of a perimarine fluviatile area. Geology and
 paleobotany of the area surrounding the archeological
 excavation at the Hazendonk river dune (Western
 Netherlands). *Analecta Praehistorica Leidensia* 16:
 1–124.
van Dijk, G. J., Berendsen, H. J. A. and Roeleveld, W.
 1991 Holocene water level development in The Neth-
 erlands' river area; implications for sea-level recon-
 struction. *Geologie en Mijnbouw* 70: 311–326.
van Geel, B. 1972 Palynology of a section from the raised

peat bog "Wietmarscher Moor", with special reference
 to fungal remains. *Acta Botanica Neerlandica* 21:
 261–284.
_____1978 A palaeoecological study of Holocene peat
 bog sections in Germany and The Netherlands, based
 on the analysis of pollen, spores and macro- and mi-
 croscopic remains of fungi, algae, cormophytes and
 animals. *Review of Palaeobotany and Palynology* 25:
 1–120.
van Geel, B. and Mook, W. G. 1989 High-resolution [14]C
 dating of organic deposits using natural atmospheric
 [14]C variations. *Radiocarbon* 31(2): 151–155.
Witte, H. J. L. and van Geel, B. 1985 Vegetational and en-
 vironmental succession and net organic production be-
 tween 4500 and 800 B.P. reconstructed from a peat de-
 posit in the western Dutch coastal area (Assendelver
 Polder). *Review of Palaeobotany and Palynology* 45:
 239–300.

APPENDIX 1. STATISTICAL BACKGROUND

We consider a ^{14}C sample as a population, in the statistical sense, of elements with different ^{14}C ages, Θ, with the probability density function $P(\Theta)$. An estimate is to be made of the "real" but unknown value, Θ_r. A laboratory measurement can be viewed as an estimate of the population mean, μ. This estimate, $\hat{\mu}$, is subject to error, and thus, is not equal to the true μ value. $\hat{\mu}$ is usually taken as an estimate of Θ_r. The total estimation error equals $\varepsilon_t = (\Theta_r - \hat{\mu})$, and can be split into two separate errors as

$$\varepsilon_t = \varepsilon_n + \varepsilon_m = (\Theta_r - \mu) + (\mu - \hat{\mu}).$$ (1.1)

Because Θ_r is unknown, it may be any of the Θ values in the population. Thus, Θ_r is replaced in Equation 1.1 by the stochas-
tic variable Θ representing all the possible ^{14}C ages in the sample. The total estimation error is made up partly of the natural
Θ variation of the population (due to the presence of elements with different ^{14}C ages), and partly of the error in estimating
the population mean. The latter is the measurement error of which the standard deviation, σ_m, is routinely given by the labo-
ratory. If this error has zero mean (unbiased estimate of μ), and is independent of the natural variation represented by $(\Theta - \mu)$,
ε_t has the following statistics

$$E[\varepsilon_t] = 0$$ (1.2a)

$$VAR[\varepsilon_t] = \sigma_t^2 = \sigma_n^2 + \sigma_m^2 \quad [^{14}C\ yr^2]$$ (1.2b)

where σ_n^2 = variance of Θ in the sample
σ_m^2 = variance of the ^{14}C measurement.

The laboratory error multiplier (Stuiver and Pearson 1986; Pearson and Stuiver 1986) is not taken into account in these for-
mulas. If the laboratory error multiplier is considerably higher than 1.0, it may contribute significantly to σ_t^2.

APPENDIX 2. RELATION BETWEEN TIME WIDTH AND σ_n

Figure 9 illustrates the relation between sample time width and natural Θ variation. Note that this relation only holds for un-
contaminated samples. Total Θ variation in a sample can be attributed to two sources: 1) variation due to different ages
within the sample resulting from sample time width; and 2) variation due to ^{14}C variations (wiggles)

$$\Theta = \Theta_{tw} + \varepsilon_w.$$ (2.1)

Figure 9 shows two curves. One is part of the most detailed calibration curve ($\sigma_n = 8$); the other is part of the $\sigma_n = 200$ cali-
bration curve, which can be assumed to be a straight line over such short intervals. The $\sigma_n = 200$ line represents the linear
trend of the calibration curve, which can be viewed as the variation of Θ due to different ^{14}C ages within the sample. The
slope α of the $\sigma_n = 200$ line is determined by:

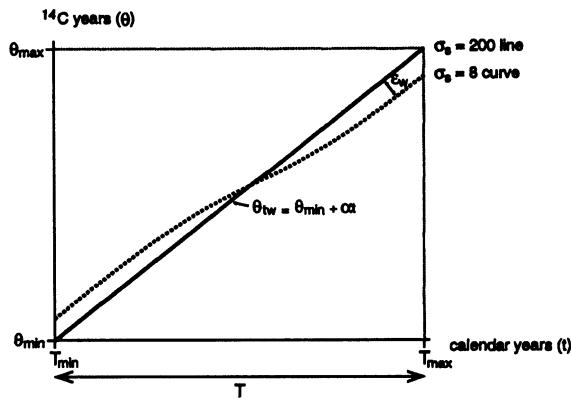$$\alpha = (T_{max} - T_{min}) / (\Theta_{max} - \Theta_{min}).$$ (2.2)

Fig. 9. Determination of parameters for the conversion of time width values to values of σ_n. Small parts of the high-resolution ($\sigma_s = 8$) calibration curve and a strongly smoothed curve without wiggles ($\sigma_s = 200$) are depicted. α is the slope of the $\sigma_s = 200$ line, which is assumed to be straight over such short intervals. σ_w is the standard deviation of the difference between the two calibration curves.

Both Θ_{tw} and ε_w are random, because the actual values are unknown. Θ_{tw} can be written as a function of time according to (see Fig. 9)

$$\Theta_{tw}(t) = \Theta_{min} + \alpha t$$ (2.3)

where the exact but unknown calendar age, t, is now the random variable. Because t has a uniform probability distribution within the time width T ($T_{max} - T_{min}$), the mean and the variance of Θ_{tw} can be calculated easily with

$$\overline{\Theta}_{tw} = \Theta_{min} + (\alpha T)/2$$ (2.4)

$$(\sigma_{\Theta tw})^2 = \alpha^2 T^2 / 12.$$ (2.5)

Hence, the natural sample error becomes

$$\sigma_n(t,T) = \sqrt{((\sigma_{\Theta tw})^2 + \sigma_w^2)} = \sqrt{(\alpha^2 T^2/12 + \sigma_w^2)} \quad [^{14}C \text{ yr}].$$ (2.6)

The slope of the linear trend of the calibration curve, α, and the standard deviation of the wiggles, σ_w, can be determined from the relevant part of the calibration curve.

APPENDIX 3. CALCULATION OF I_i^{obs} AND I_i^{exp}

The observed intensity, I_i^{obs}, for each C-class is obtained by dividing the intensity for each C-class, I_i^c, by the total intensity of the "real" ^{14}C histogram, I^{tot} (Fig. 10A)

$$I_i^{obs} = I_i^c / I^{tot}$$ (3.1)

where $\sum_{i=1}^{8} I_i^{obs} = 1.$

If the large-scale distribution of ^{14}C ages in the "real" ^{14}C histogram is homogeneous, $I_i^{exp} = 0.125$. However, due to the irregular distribution of the C-classes over the ^{14}C time scale (Figs. 3C–E) and the unequal large-scale distribution of ^{14}C ages in the "real" ^{14}C histograms (Fig. 1), this is not the case, making it necessary to calculate a weighted expected intensity (I_i^{exp}) for each C-class. The ^{14}C axis of the "real" ^{14}C histograms is subdivided into 12 segments of 500 ^{14}C yr each (110–600, 610–
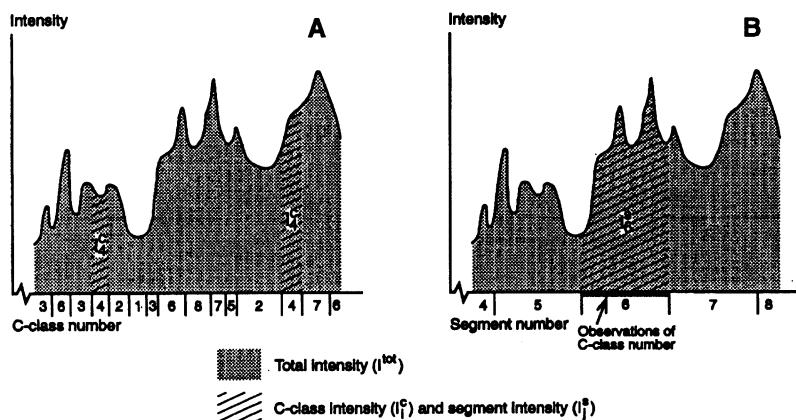
Fig. 10. Determination of total intensity, I^{tot}, (A) intensity of individual C-classes, I_i^c, (B) intensity of individual segments, I_j^s

1110, ..., 5610–6100 yr BP), and the intensity for each segment, I_j^s, is determined (Fig. 10B). The weight factor for each segment is

$$\lambda_j = I_j^s / I^{tot} \tag{3.2}$$

where $\quad \sum_{j=1}^{12} \lambda_j = 1.$

In each segment, the C-class number has been determined for every 10th ^{14}C yr, totaling 50 observations per segment (Fig. 10B). The number of times that C-class i is encountered in segment j, C_{ij}, is determined, and the relative frequency of each C-class for each segment is obtained by

$$RF_{ij} = C_{ij}/50 \tag{3.3}$$

where $\quad \sum_{i=1}^{8} RF_{ij} = 1.$

The expected intensity for each C-class is then calculated as

$$I_i^{exp} = \sum_{j=1}^{12} (\lambda_j \cdot RF_{ij}) \tag{3.4}$$

where $\quad \sum_{i=1}^{8} \left(\sum_{j=1}^{12} (\lambda_j \cdot RF_{ij}) \right) = 1.$

APPENDIX 4. RELATION BETWEEN σ_s AND σ_t

Because the smoothing parameter, σ_s, is expressed in cal yr and the total sample error, σ_t, in ^{14}C yr, their mutual relation is not straightforward. This can be problematic in cases where σ_s-a plots cannot be calculated (*e.g.*, due to insufficient data), and the choice of a calibration curve is based on an estimate of σ_t. σ_t is a mean value for large data sets that are nearly always distributed over several millennia. Because σ_s is constant for the whole time scale, a correction factor, Q, can be determined for converting σ_t into σ_s by

$$Q = \Delta \text{ cal yr} / \Delta \,^{14}C \text{ yr} . \tag{4.1}$$

In this equation, $\Delta \,^{14}C$ yr comprises the range of the total data set, provided that the ^{14}C ages are roughly equally distributed in time. Calibration of the upper and lower limit of the ^{14}C data range with the Groningen program (van der Plicht and Mook 1989; van der Plicht 1993), using a $\sigma_s = 200$ smoothed calibration curve yields Δ cal yr.

LIST OF SYMBOLS

a	slope of the I_i^{obs}/I_i^{exp} ratio regression line
C	parameter representing the degree of clustering of ^{14}C ages due to ^{14}C variations
C_{ij}	number of occurrences of ith C-class in jth segment
I	intensity of a ^{14}C histogram
I_i^c	intensity of ith C-class
I_i^{exp}	expected intensity of ith C-class in a ^{14}C histogram
I_i^{obs}	observed intensity of ith C-class in a ^{14}C histogram
I_j^s	intensity of jth segment
I^{tot}	total intensity of a ^{14}C histogram
Q	correction factor for determining the relation between σ_t and σ_s
RF_{ij}	relative frequency of ith C-class in jth segment
t	real age of a ^{14}C sample [cal yr]
T	time width of a ^{14}C sample [cal yr]
α	slope of the linear trend of part of the $\sigma_s = 200$ calibration curve
ε_m	measurement error
ε_n	natural sample error
ε_t	total estimation error of the true ^{14}C age
ε_w	error due to variation caused by wiggles
Θ	^{14}C age of an individual element of a ^{14}C sample
Θ_r	real age of a ^{14}C sample [^{14}C yr]
Θ_{tw}	variation of ^{14}C ages caused by time width [^{14}C yr]
λ_j	weight factor
μ	mean age of a ^{14}C sample [^{14}C yr]
$\hat{\mu}$	estimation of mean age of a ^{14}C sample [^{14}C yr]
σ_m	standard deviation of measurement error [^{14}C yr]
σ_n	standard deviation of natural sample error [^{14}C yr]
σ_s	smoothing parameter [cal yr]
σ_t	standard deviation of total sample error [^{14}C yr]
σ_w	standard deviation of wiggles *vs.* a regression line [^{14}C yr]
$\sigma_{\Theta tw}$	standard deviation of variation of ^{14}C ages caused by time width [^{14}C yr]