

CHECKING BACK ON AN ASSEMBLAGE OF PUBLISHED RADIOCARBON DATES

M G L BAILLIE

Palaeoecology Centre, The Queen's University, Belfast BT7 1NN, Northern Ireland

ABSTRACT. It is clear that radiocarbon researchers take a forward view towards the improvement of accuracy and precision in dating. Unfortunately, archaeologists base much of their research on the published dates produced in the past. Archaeologists and other users of radiocarbon dates should understand the limitations associated with past dates. This article addresses these limitations by looking at a large number of routine radiocarbon dates associated with a block of English tree-ring chronologies, the true ages of which are now known within close limits. My conclusion supports the idea of global multiplication factors as proposed by the International Study Group (1982).

INTRODUCTION

The period 1985 - 1988 saw an intensive effort to outline a prehistoric oak tree-ring chronology in England (Baillie & Brown 1988). This work centered on sub-fossil oaks from East Anglia and Lancashire and built on a previous chronology from Swan Carr, near Durham, which spanned 381 - 1155 BC (Baillie, Pilcher & Pearson 1983). The approach to chronology building was to produce well-replicated chronology units which could be located precisely in time against the existing Irish (Pilcher *et al* 1984) and North German (Leuschner & Delorme 1984) chronologies. This dating proved highly successful and it became apparent that a high percentage of the East Anglian oaks fell in the calendar period 3200 - 1680 BC, whereas Lancashire oaks spanned a much greater range. Lancashire chronologies ultimately spanned 970 - 1584 BC, 1681 - 3807 BC and 4165 - 4989 BC. The 17th century BC gap was resolved using the Hasholme 1326 - 1687 BC bog-oak chronology supplied by J Hillam at Sheffield.

Our essentially random sampling of Lancashire bog oaks failed to identify any material across the period 3807 - 4165 BC. The English Neolithic complex of three chronologies from the River Trent, from the Sweet Track and from a submerged forest at Stolford (Morgan, Litton & Salisbury 1987) offered a possible solution to this problem. This chronology, which spanned 631 years as published, proved difficult to tie down against the Irish and German chronologies. As Hillam (1988) pointed out, this failure makes for "one of the major puzzles in British dendrochronology."

PLACING THE NEOLITHIC CHRONOLOGY

We acquired fresh samples from Stolford and the Nottingham Laboratory supplied the original River Trent samples for remeasurement. It quickly became apparent that the River Trent 576-year chronology (Salisbury *et al* 1984) was, in fact, made up of two chronologies. One, Old Loop 1, spanned 354 years and cross-dated with Sweet and Stolford as published (Morgan *et al* 1987). The other, Colwick Hall 1, spanned 349 years and was placed by a tree-ring match at 2697-3045 BC against both the East Anglian chronology ($t = 6.1$) and the Lancashire chronology ($t = 5.0$). A new radiocarbon date (UB-3055, 4167 ± 40 BP) confirmed this new placement.

Our interest in this splitting of the River Trent chronology was that the *original* linking of the Old Loop and Colwick Hall chronologies had been at least partly conditioned by two radiocarbon determinations, 5110 ± 45 BP and 5335 ± 50 BP, respectively. It was now apparent that the 5335 ± 50 BP date was in error by > 1000 ^{14}C years! I do not intend to go into a lengthy discussion of this "bad" date. It could be a laboratory measurement error; it could be a sample mix-up either on the part of the dendrochronologists or the laboratory; it could be a computational error or a simple number transcription. It would not have been possible to isolate the specific cause at the time, nor is it now. The fact is that an erroneous date went into the literature, unchallenged, and

was, at least in part, responsible for an incorrect tree-ring linkage. It is also worth reflecting that this was a high-integrity wood sample which did not suffer from the sampling vagaries of most archaeological samples.

As research continued into the dating of the reduced 438-year Neolithic chronology, there was a further revelation. We discovered that the Stolford timber, which represented the outer 272 years of the complex, had been used in the first Glasgow interlaboratory comparison exercise (International Study Group 1982). Eight samples from this timber had each been dated by up to 20 laboratories, some at high precision. The resulting weighted mean dates (oldest to youngest) were 5168, 5196, 5173, 5094, 5070, 5031, 5025 and 5030 BP. All had standard error of the mean values of < 10 years and could thus be classed as high-precision dates. We could then "wiggle match" this component of the Neolithic chronology against the high-precision calibration curve of Pearson *et al* (1986). This showed that the end of the 410-year Sweet Track chronology was likely to fall within ca 10 years of the calendar date 3800 BC (Baillie & Pilcher 1988).

With this new information, we could look at two packages of routine radiocarbon dates associated with the Sweet Track, itself. First, there were 11 dates on the wood from the track (Morgan 1988) (Fig 1). A high proportion of the dates are too young. Even allowing for the original laboratory disclaimer on one replicate sample (which was ca 1000 ^{14}C years too young) and the general statement about root and insect contamination, these dates remain in the literature. In fact, Clark and Morgan used these dates in an earlier wiggle-matching exercise. Their suggested end date for the Sweet track was in the range 3685 - 3415 BC, with 95% confidence (Morgan 1988). It is now clear that this range does not include the correct date and the center point of the range is more than two centuries too young.

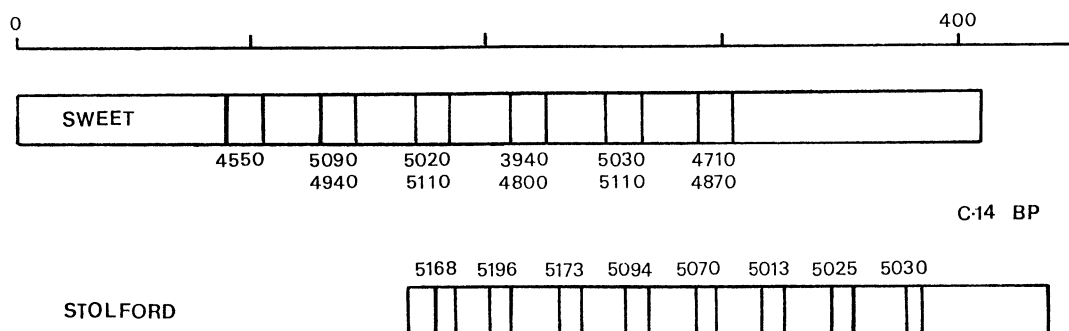


Fig 1. The high-precision results for Stolford, from the First Glasgow Interlaboratory Study, compared with the routine radiocarbon dates for the Sweet Track

In addition to the dates on the wood of the track itself, are ten samples on peat and stakes associated with the track (Morgan 1988). If the track was truly constructed within a few years of 3800 BC, then reference to the calibration curve shows that *all samples* younger (closer to the present) than the track should give *radiocarbon ages* younger than 5000 BP, whereas all samples older (further back in time) than the track should give radiocarbon ages older than 5000 BP.

Table 1 shows the actual results. The dates for peat older than and substantially older than the track are particularly revealing as they are much too young and suggest either poor pretreatment or a significant laboratory bias. Suggestions that peat is a notoriously difficult medium for dating hardly encourage the archaeologist, as there was no prior warning that the samples might give erroneous dates. It hardly befits radiocarbon laboratories to discount dates that were regarded as perfectly adequate in a "blind dating" exercise, *ie*, where the customer did not know the correct answer.

TABLE 1
Radiocarbon ages for the Sweet Track

	Samples directly associated with the Sweet Track	Radiocarbon age BP
1	Hazel peg 325	5218 ± 75
2	Hazel slat	5159 ± 70
3	Hazel peg 303	5150 ± 65
4	Hazel and ash pegs	4887 ± 90
5	Peat beneath rails	5140 ± 100
6	Peat packing	5103 ± 100
7	Peat on plank	5224 ± 75
8	Peat on plank	5108 ± 65
9	Peat below track	4744 ± 45
10	Peat well below track	4848 ± 45

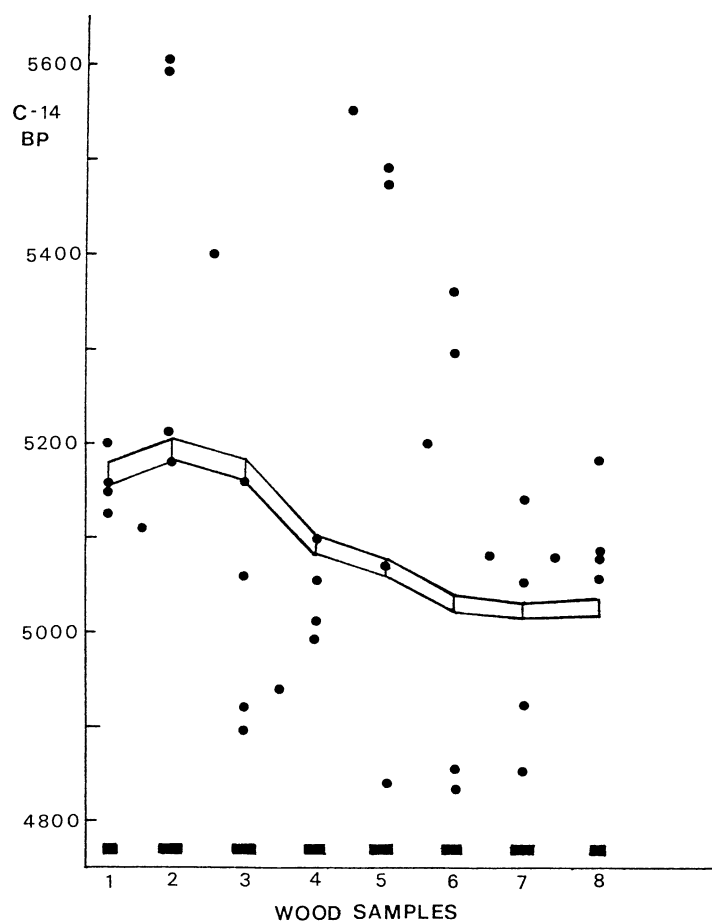


Fig 2. Routine dates (●) on samples from the Stolford timber compared with the high-precision results from the First Glasgow Interlaboratory Study (□)

Finally, we also noted that the Glasgow Laboratory had used samples from the Stolford timber to test reproducibility on their own radiocarbon system. Again, the results should not be used to judge the present laboratory, as they were used as a stepping stone towards improving the accuracy of the system. However, the results form a perfectly good guide to the general performance of a typical routine laboratory up to the early 1980s and are included here as such.

Some 39 samples can be identified (Scott, Baxter & Aitchison 1983) that relate directly to the samples used in the International Interlaboratory Study. Figure 2 shows the distribution of the center points of the date ranges compared with the high-precision results cited above. Clearly, the results show significant random variation with an apparent bias towards "too old" results.

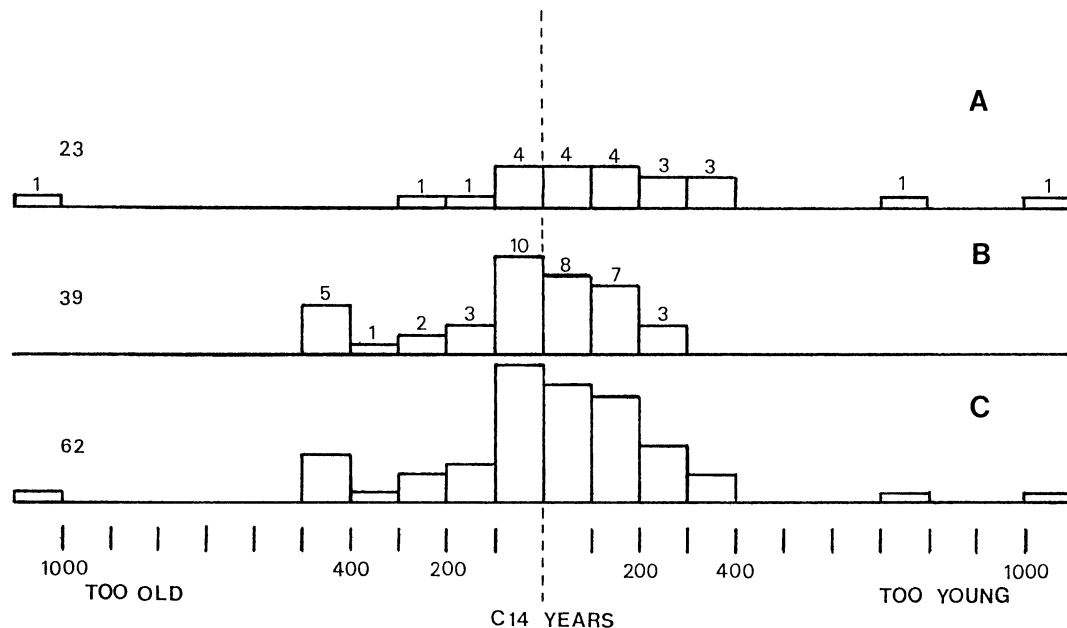


Fig 3. Frequency of departure of the center points of routine dates from their true radiocarbon ages. A - for the 23 dates associated with the River Trent and the Sweet Track; B - for the 39 dates associated with the Stolford timber; C - all 62 routine dates associated with the Neolithic complex.

CONCLUSION

Our study uncovered some 62 routine radiocarbon determinations associated with the English Neolithic chronology complex. Figure 3 combines these dates, showing the frequency of departure of the center points of the dates from their true radiocarbon ages. The simplest analysis of these results appears in Table 2.

TABLE 2
Percentages of dates with center points greater than
the stated distances from their true radiocarbon ages

58% lie outside 100 radiocarbon years
34% lie outside 200 radiocarbon years
19% lie outside 300 radiocarbon years
13% lie outside 400 radiocarbon years
5% lie outside 500 radiocarbon years

Of these figures, the most damning for routine radiocarbon analysis is the 34% outside 200 radiocarbon years. From this, it would seem that a full one-third of all radiocarbon dates are effectively useless from the point of view of tight chronological research. We need to give some thought to the implications of these results. Here are some of my views:

1. The quoted errors are highly unreliable in about one-third of all samples.
2. Dates on dubious samples, particularly those with possible contamination, be it mobile humic acid or rootlets or both, are allowed to enter the literature with little comment.
3. The blind nature of most dating procedures, when the customer has no idea of the true age, means that bad dates are neither suspected nor easily detected.
4. These 62 dates are not atypical. They should be a perfectly typical sample of all the routine dates in the literature.
5. 95% confidence, as applied in the case of Clark and Morgan's wiggle match (Morgan 1988) of the Sweet dates, means nothing if the dates are wrong or systematically biased.
6. On the positive side, no date was significantly more than one millennium in error.

I have no intention to embarrass the labs concerned. Undoubtedly they are trying to improve and, indeed, things may have improved in the meantime. These results should not be taken as necessarily representative of the results of these laboratories in general. However, all things considered, I see no reason to suppose that this whole assemblage of dates is unrepresentative of the vast bulk of archaeological dates in the literature. If this is the case, then the figures in Table 1 represent a basis for assessing the past ability of a wide range of radiocarbon laboratories and should be taken into account when archaeologists attempt to infer absolute dates from radiocarbon dates.

In the final analysis, routine radiocarbon dates do not conform to the same standards of absolute dating as dendrochronology. We must use great care in attempting to pinpoint any event in real time with routine radiocarbon dates - even where large numbers of routine radiocarbon dates are involved. However, some comfort can be drawn from the fact that, according to this study, *only* 5% are outside 500 radiocarbon years from the true radiocarbon age.

In summary, radiocarbon dating is a very useful method for putting most samples into the right chronological ball park. Used as a fairly blunt tool, it is fine most of the time - hence its success. Unfortunately, radiocarbon researchers have allowed themselves the belief that their results are within the statistical limits quoted, largely because most of their customers do not complain about the dates (because the customer does not know the true date). Most laboratories calibrate their dates using the quoted errors - they do not habitually use multiplication factors to allow for additional errors.

From this 62-date test sample, a significant proportion of routine dates are *not* within the statistical limits quoted. As such, they are not useful for resolving highly detailed chronological questions. It is also interesting that the global-average-quoted error on the 62 dates happens to be ± 79 years. If we used the multiplication factors suggested by Scott, Baxter and Aitchison (1983),

dates of similar quality might well have their quoted errors multiplied by 2.6. In other words, these 62 routine dates should have had quoted errors of $\text{ca } \pm 200$ years. If this were applied globally to the 62 dates, it would account for the spread of some 90% of the samples (Fig 3). This suggests strong support for *multiplication factors*, a concept that so far has found little favor with archaeologists but would add some realism to routine dating.

This study leads to the conclusion that laboratories should not make extravagant claims about the accuracy and precision of their dates unless they have *demonstrated proof* for such claims.

ACKNOWLEDGMENTS

The author wishes to thank the Nottingham Tree-Ring Group for supplying original timbers for re-examination and J Hillam for important raw data. The new radiocarbon determination on the Colwick Hall timber was kindly supplied by Dr GW Pearson.

REFERENCES

- Baillie, MGL and Brown, DM 1988 An overview of oak chronologies. *Br Archaeol Repts*, Br ser 196: 543-548.
- Baillie, MGL and Pilcher, JR 1988 Make a date with a tree. *New Scientist* 117: 48-51.
- Baillie, MGL, Pilcher, JR and Pearson, GW 1983 Dendrochronology at Belfast as a background to high-precision calibration. In Stuiver, M and Kra, RS, eds, Internatl ^{14}C conf, 11th, Proc. *Radiocarbon* 25(2): 171-178.
- Hillam, J 1988 Dendrochronology - 20 years on. *Current Archaeology* 107: 358-363.
- Hillam, J, Groves, CM, Brown, DM, Baillie, MGL, Coles, JM and Coles, BJ 1990 Dendrochronology of the English Neolithic. *Antiquity* 64: 210-220.
- International Study Group 1982 An inter-laboratory comparison of radiocarbon measurements in tree-rings. *Nature* 298: 619-623.
- Leuschner, HH, von and Delorme, A 1984 Verlängerung der Gottingen Eichenjahrringchronologien für Nord- und Süddeutschland bis zum Jahr 4008 v Chr. *Forstarchiv* 55: 1-4.
- Morgan, RA 1988 Tree-ring studies of wood used in Neolithic and Bronze Age trackways from the Somerset Levels. *Br Archaeol Repts* Br ser 184:
- Morgan, RA, Litton, CD and Salisbury, CR 1987 Trackways and tree trunks - Dating Neolithic oaks in the British Isles. *Tree-Ring Bull* 47: 61-69.
- Pearson, GW 1986 Precise calendrical dating of known growth-period samples using a 'curve fitting' technique. In Stuiver, M and Kra, RS, eds, Internatl ^{14}C conf, 12th, Proc. *Radiocarbon* 28(2A): 292-299.
- Pilcher, JR, Baillie, MGL, Schmidt, B and Becker, B 1984 A 7272-year tree-ring chronology for western Europe. *Nature* 312: 150-152.
- Salisbury, CR, Whitley, PJ, Litton, CD and Fox, JL 1984 Flandrian courses of the River Trent at Colwick, Nottingham. *The Mercian Geologist* 9(4): 189-207.
- Scott, EM, Baxter, MS and Aitchison, TC 1983 ^{14}C dating reproducibility: Evidence from a combined experimental and statistical programme. In Waterbolk, HT and Mook, WG, eds, Internatl symposium ^{14}C and Archaeology, Proc. *PACT* 8 (II.7): 133-145.

Note added in proof: Since this article was written, fresh evidence has allowed the definitive dating of the Sweet Track chronology. The oak timbers used in the construction of the Sweet Track last grew in 3807 BC and were felled either late in 3807 or early in 3806 BC (Hillam *et al* 1990). This date is extremely close to that predicted by high-precision radiocarbon analysis (see text).