

SECTION 7: CHARACTERIZATION OF THE REFERENCE MATERIALS BY CONSENSUS VALUES

7.1 CALCULATION OF CONSENSUS VALUES

Each material needs to be characterized by estimating its activity, which creates a reference value for each material. This value then can be considered as the "known" activity of the material and future analyses can be compared to this to quantify the accuracy of the measurement. In this way, the materials remain useful for laboratory quality assurance.

The procedure used in the calculation of the consensus value comes from Rozanski et al. (1992) and is an iterative one. It is described below.

There are 3 stages.

- **Stage 1**: Outlying results are removed if they are greater than 3 interquartile ranges from the nearest of either the lower or upper quartiles. This occurs when a result is either greater than Q3 + 3(Q3–Q1) or less than Q1 – 3(Q3–Q1), where *Q1* and *Q3* are the lower and upper quartiles, respectively. Then, the preliminary consensus value is calculated as the median (*m*) of the remaining results.
- **Stage 2**: Remove results that are at least twice their quoted error (σ) from the preliminary consensus value. That is, only keep $|x–m| / \sigma < 2$, where *x* is the result, *m* the preliminary consensus value, and σ the quoted error.
- **Stage 3**: Calculate the final consensus as a weighted mean of the remaining results, using their σ^2 values as the weights.

7.1 Remarks on the Procedure

For FIRI Samples A and B (in yr), this approach is not very appropriate given that many laboratories did not quote finite ages. For these samples, an alternative approach was used based on the reliability analysis (see Section 6).

It should also be noted that averages are rather sensitive to extreme data values, which is why the outliers are removed in Step 1.

The approach has the advantage that an estimated error can be calculated for the consensus value (which will usually be very small since there are a large number of results).

7.2 INITIAL CONSENSUS VALUES

Consensus values are reported in Table 7.1, based on the Rozanski et al. procedure.

Table 7.1 Preliminary consensus values

FIRI sample	Weighted mean (1 σ)	AMS	GPC	LSC
C	18,173 (10.5)	18,183 (13)	18,229 (28)	18,140 (25)
DF	4508 (3)	4519 (4)	4484 (5)	4507 (6)
E	11,778 (7)	11,805 (9)	11,738 (19)	11,707 (17)
GJ	110.69 (0.04)	110.52 (0.05)	110.85 (0.07)	110.82 (0.08)
H	2232 (5)	2238 (6)	2198 (9)	2233 (9)
I	4485 (5)	4483 (7)	4456 (10)	4499 (11)

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

It should be noted that the results for Samples A and B are not included in this table, since this procedure only is possible using results where a quoted error is given. However, the results for Samples A and B will be returned to later in this section, when the analysis of the pMC is completed.

Figures 7.1 to 7.7 (Section 7 appendix, p 269–275) show the distribution of the laboratory results around these consensus values. They include the laboratory-quoted errors. In such figures, we can see how closely the results from the different laboratories agree (accounting for their quoted errors). The consensus values are also marked. In Step 2, laboratories quoting small errors will be excluded, unless they lie close to the consensus value, while laboratories quoting large errors will be included in Step 3. However, in Step 3, results with large errors will be down-weighted in the calculation and so will not have a large impact on the final result.

Therefore, there is an issue of how robust the initial consensus value is in Step 1, and how important its definition is on the final consensus value. Therefore, we consider variants of this original method, which at Step 1 exclude not simply extreme age/activity values, but also results with large quoted errors.

7.3 THE EFFECT OF SCREENING OUT RESULTS WITH LARGE QUOTED ERRORS IN CONSENSUS CALCULATIONS

7.3.1 σ Method 1

This method is the same as the original one, except that, between Stage 1 and Stage 2, results with a quoted uncertainty greater than a certain cut-off point are rejected.

7.3.2 σ Method 2

This method is the same as the σ Method 1, but this time, results with a quoted uncertainty greater than a certain cut-off point are rejected *before* Stage 1.

7.3.3 Choice of σ "Cut-off" Points

The choice of the cut-off points is subjective. However, from the histograms showing the distribution of σ and expert opinion, the cut-off points shown in Table 7.2 were used for both methods. Because of the subjectivity of the decisions, 2 (or, in AB's case, 3) different values for the cut-off points were chosen for each sample.

Table 7.2 Cut-off points used for the different samples

Sample	Cut-off points used			Units
Kauri wood, AB	0.3	0.15	0.1	pMC
Turbidite, C	200	150	—	yr BP
Belfast dendro-wood, DF	100	50	—	yr BP
Humic acid, E	150	100	—	yr BP
Barley mash, GJ	1	0.6	—	pMC
Hohenheim dendro-wood, H	100	50	—	yr BP
Belfast cellulose, I	100	50	—	yr BP

7.3.4 Results

From the results in Table 7.3, it can be seen that the various methods for calculating the consensus make very little difference to all but the AB sample (ranges of only 2.3 yr BP for Sample C, 1.4 yr BP for DF, 11.7 yr BP for Sample E, 0.06 pMC for GJ, 0.7 yr BP for H, and 10.6 yr BP for I).

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

For Sample AB, there is little difference within the σ Method 1 (a range of only 0.075 pMC), but there is for the σ Method 2, with the more restrictive cut-off points. When results with a σ greater than 0.1 pMC are screened, the consensus value becomes 0.2 pMC, less than two-thirds the value it is under the original method.

Table 7.3 Consensus values under the different methods

Sample		Original methods	Method 1			Method 2		
AB	σ cut-off	None	0.3	0.15	0.1	0.3	0.15	0.1
(pMC)	Consensus	0.330	0.330 (0.01)	0.327 (0.01)	0.325 (0.01)	0.324 (0.01)	0.251 (0.01)	0.203 (0.01)
C	σ cut-off	None	200	150	—	200	150	—
(yr BP)	Consensus	18,175.5 (10.5)	18,176.5 (9.7)	18,177.8 (9.3)	—	18,176.5 (9.7)	18,177.8 (9.3)	—
DF	σ cut-off	None	100	50	—	100	50	—
(yr BP)	Consensus	4508.3 (3)	4508.2 (3)	4506.8 (3)	—	4508.2 (3)	4506.8 (3)	—
E	σ cut-off	None	150	100	—	150	100.00	—
(yr BP)	Consensus	11,779.9 (7)	11,781.2 (8)	11,781.7 (7.6)	—	11,791.6 (7.8)	11,791.2 (8)	—
GJ	σ cut-off	None	1	0.6	—	1	0.6	—
(pMC)	Consensus	110.69 (0.04)	110.69 (0.04)	110.72 (0.04)	—	110.69 (0.04)	110.75 (0.04)	—
H	σ cut-off	None	100	50	—	100	50	—
(yr BP)	Consensus	2232.5 (5)	2232.3 (4.7)	2233.0 (4.7)	—	2232.3 (4.7)	2233.0 (4.7)	—
I	σ cut-off	None	100	50	—	100	50	—
(yr BP)	Consensus	4484.9 (5)	4485.1 (5)	4482.1 (5)	—	4485.1 (5)	4474.5 (5.3)	—

7.3.5 Discussion

The alternative methods for calculating the consensus only lead to very small differences, except in the case of the Kauri wood sample, AB. Here, screening out results with σs larger than a cut-off point before using the original method, shifted the consensus by large amounts when the cut-off was small (from 0.33–0.20 pMC, when the cut-off was 0.1 pMC). Possible reasons for this change could stem from AB being a sample at, or near, the limits of detection for ^{14}C dating.

Since the Kauri wood's activity is so low, some results are given as background or non-finite. This occurs when the σ is large with respect to its result. Obviously, those laboratories that have a lower σ can give finite results for older samples. Because background and non-finite results are excluded from the consensus calculation, this could bias the calculations.

Also, it is possible that laboratories have reported pMC results for samples that should be considered background or non-finite. At present, these results are not screened out. Such an approach could be valuable in providing a more reliable estimate of the activity in the Kauri wood sample.

7.3.6 Conclusion

The consensus calculations are robust in the initial screening stages for all but the Kauri wood samples. For this sample, the consensus age has been calculated by a different method and reported in Section 6. For the pMC results, the consensus calculation has been carried out, but with a number

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

of caveats. In AB's case, a better value for the consensus may be achieved if any results that are too small with respect to their quoted errors are screened out. If this does not help, then we may be left with a large range for the consensus value of the Kauri wood sample's activity. We recommend that the results in Table 7.1 and Table 6.20 be used as consensus values for the FIRI samples.

7.4 DEVIATIONS FROM CONSENSUS VALUES

We define the standardized deviation as the difference between the result and the consensus value, divided by the quoted uncertainty on the result. Using this summary, we can explore the distribution of laboratory performance. Ideally, we might expect a standardized deviation to lie between +2 and –2. Values greatly exceeding 2 or –2 indicate either a large absolute difference between the result and the consensus value or a "large" difference relative to the quoted error. This makes them sensitive indicators of general laboratory performance. The standardized deviations for each sample (except AB) can then be investigated for the effects of different laboratory factors.

7.4.1 Effect of Laboratory Type for Sample C: Turbidite

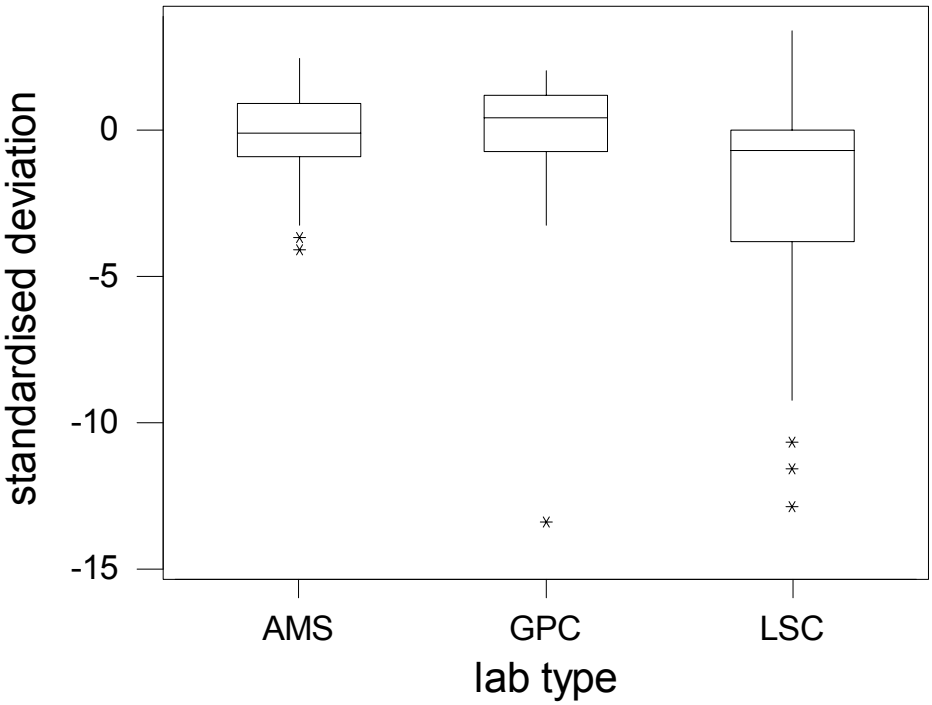

Figure 7.8 Distribution of standardized deviation for Sample C

We can see that AMS and GPC results appear to show a broadly similar distribution. For LSC results, the distribution is more widely scattered. Each laboratory type has a number of extreme values and this is more pronounced for the LSC set of results.

7.4.2 Effect of Laboratory Type for Sample D: Belfast Wood

A similar pattern is apparent; the median value lies close to 0, but there are a number of extreme values, typically reported by LSC laboratories.

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

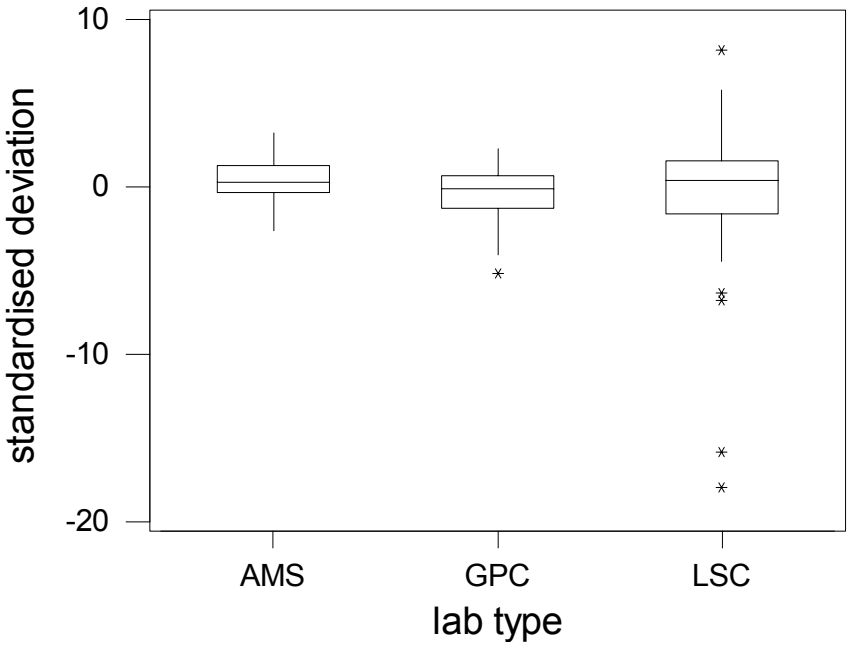

Figure 7.9 Distribution of standardized deviation for Sample D

7.4.3 Effect of Laboratory Type for Sample E: Humic Acid

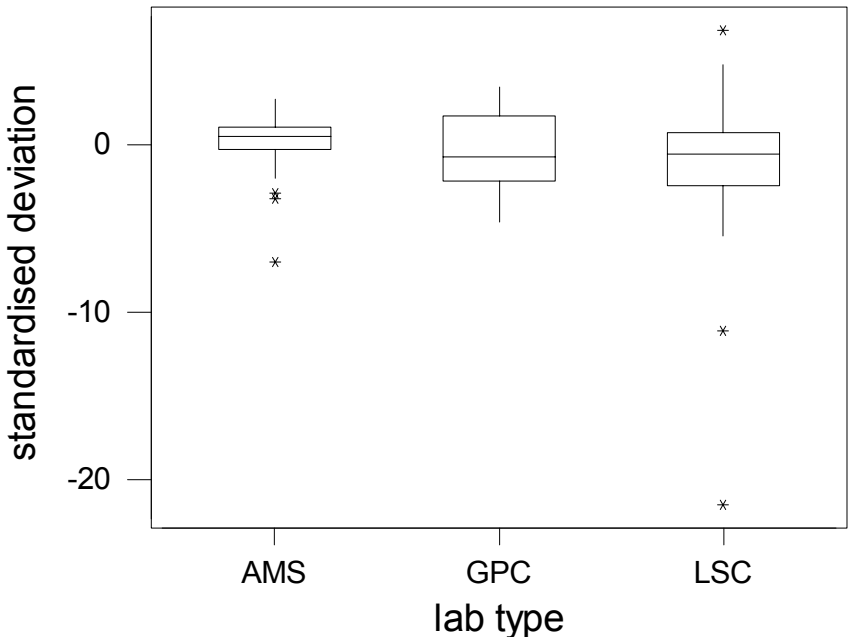

Figure 7.10 Distribution of standardized deviation for Sample E

A similar distributional pattern is apparent; the median value lies close to 0, but there are a small number of extreme values.

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

7.4.4 Effect of Laboratory Type for Sample F: Belfast Wood

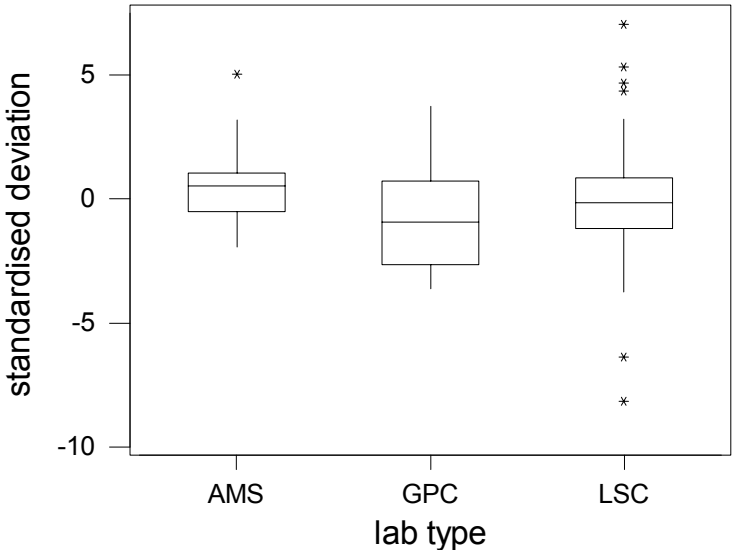

Figure 7.11 Distribution of standardized deviation for Sample F

The median value lies close to 0, but there are a number of extreme values, typically reported by LSC laboratories.

7.4.5 Effect of Laboratory Type for Sample G: Barley Mash

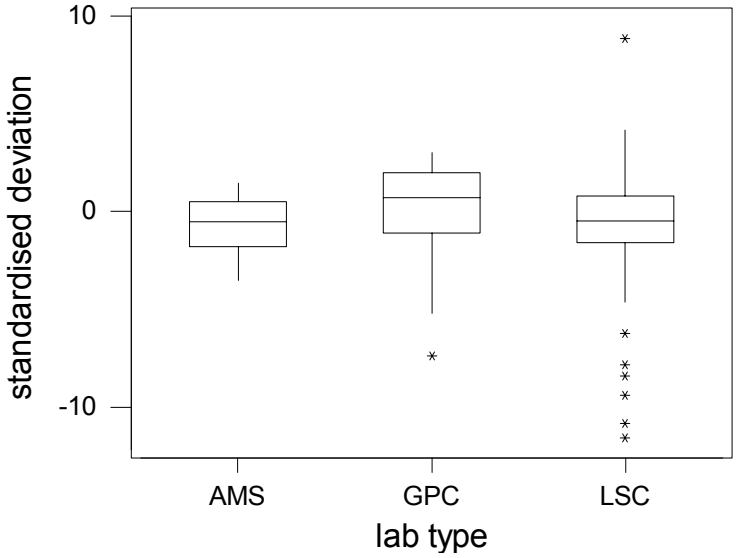

Figure 7.12 Distribution of standardized deviations for Sample G

The median value lies close to 0, but there are a number of extreme values, typically reported by LSC laboratories. Omitting these results would result in broadly similar distributions for the 3 laboratory types.

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

7.4.6 Effect of Laboratory Type for Sample H: Hohenheim Wood

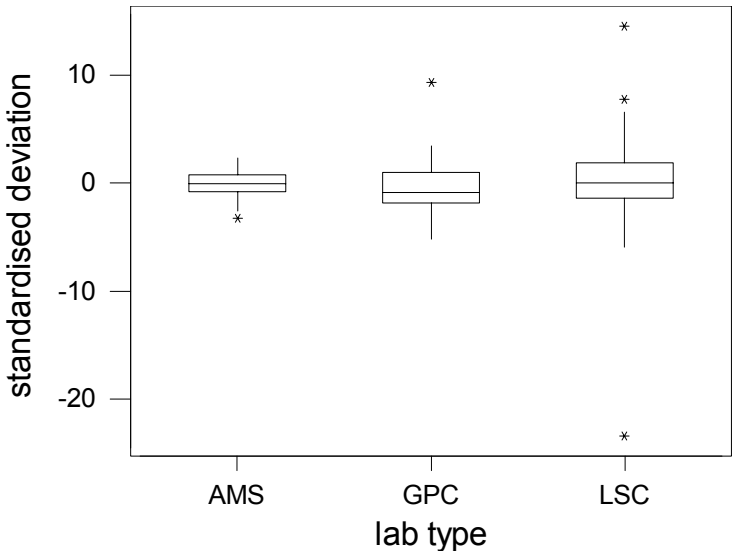

Figure 7.13 Distribution of standardized deviations for Sample H

A similar pattern is apparent. The median value lies close to 0, but there are a number of extreme values, typically reported by LSC laboratories.

7.4.7 Effect of Laboratory Type for Sample I: Belfast Cellulose

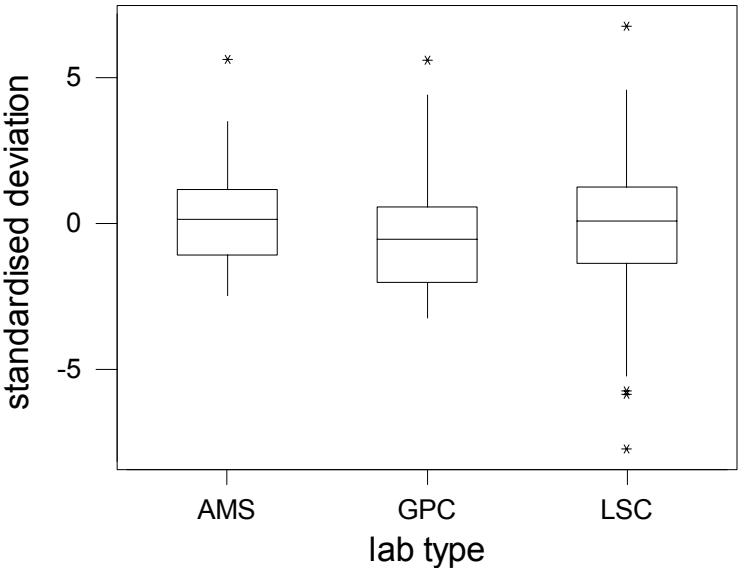

Figure 7.14 Distribution of standardized deviations for Sample I

The distribution of results is less wide for this sample. The median value lies close to 0, but there are a small number of extreme values, which are reported by LSC laboratories.

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

7.4.8 Effect of Laboratory Type for Sample J: Barley Mash

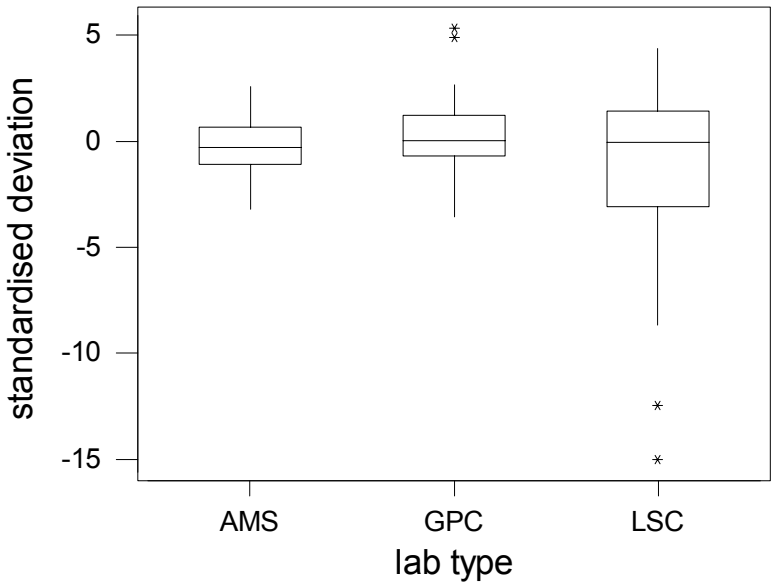

Figure 7.15 Distribution of standardized deviations for Sample J

A similar pattern is apparent, where the median value lies close to 0. The distribution of results is wider for LSC laboratories and there are several extreme values.

7.4.9 Effect of Laboratory Type for Sample A: Kauri Wood

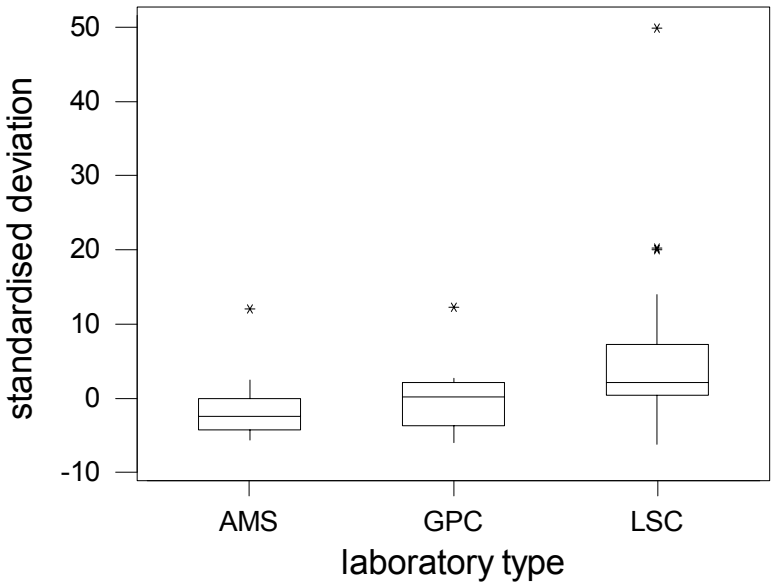

Figure 7.16 Distribution of standardized deviations for Sample A

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

7.4.10 Sample B: Kauri Wood

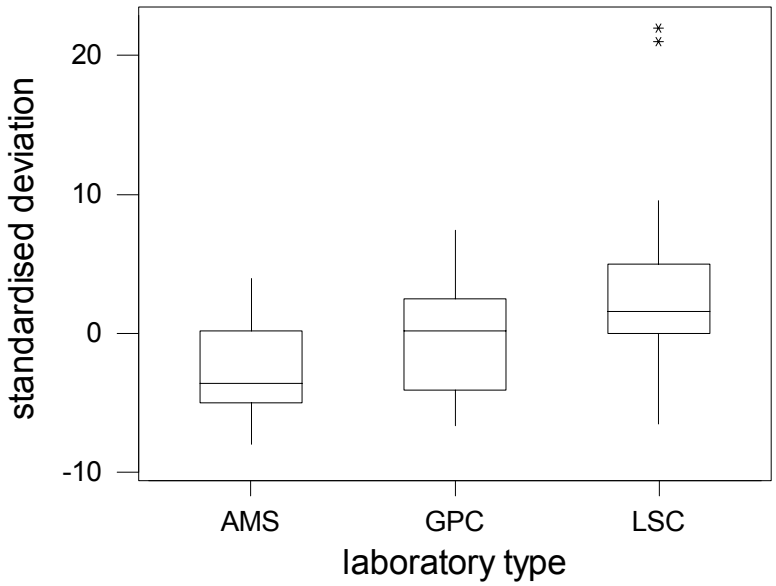

Figure 7.17 Distribution of standardized deviations for B

For Samples A and B, the calculations have been performed in pMC. We can see in Figures 7.16 and 7.17 that the distribution of results is skewed towards positive values, indicating that the laboratories reported results higher than the consensus value.

7.4.11 Effects of Other Laboratory Factors

It is of interest to explore the deviations from consensus values and to consider which factors, if any, can explain this variation. We have used the "initial" consensus values for this analysis and have not used Samples A and B. The consensus values were also all expressed in pMC to facilitate a global analysis over all the sample materials. We first consider the laboratory throughput.

The are 4 levels for the "number of analyses performed":

- *1* indicates <100 analyses done per yr by that laboratory;
- *2* indicates 100–200;
- *3* indicates 200–500;
- *4* indicates >500.

Table 7.4 Descriptive statistics for the standardized deviation by number of analyses

Nr of analyses	N	Mean	Median	StDev	Q1	Q3	Min	Max
1	109	–0.366	–0.163	4.044	–1.753	0.635	–18.15	20.25
2	266	0.753	0.380	5.156	–0.943	2.092	–15.01	49.94
3	118	–0.645	–0.040	3.770	–1.429	1.089	–19.75	11.00
4	384	–0.060	–0.202	2.537	–1.341	0.869	–8.00	12.30
Unknown	115	0.540	0.261	4.103	–0.967	1.556	–11.59	22.35

From the table, there are clearly some rather extreme values, but the IQR (Q1 to Q3) lies comfortably in the –2 to +2 range.

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

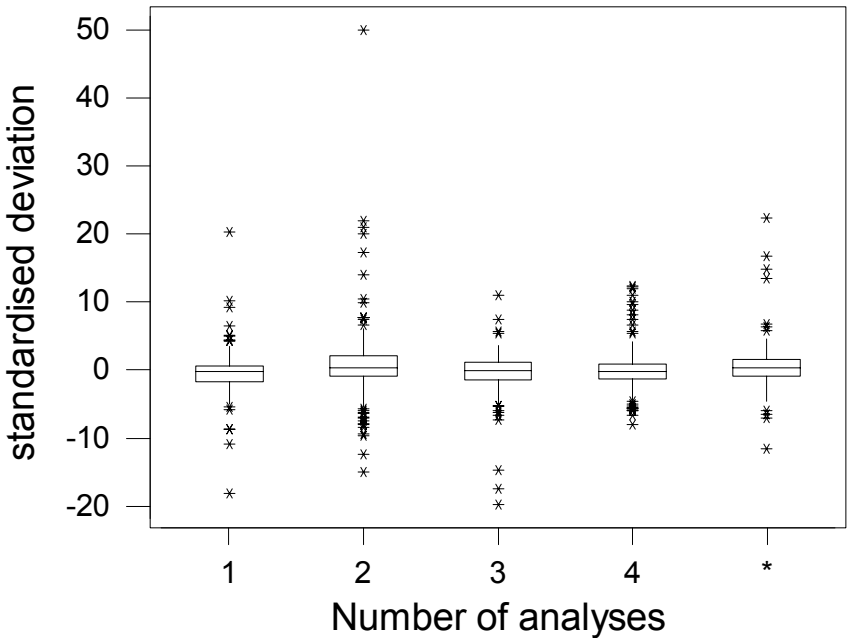

Figure 7.18 Distribution of standardized deviations by number of laboratory analyses

The results are highly skewed with many outliers. For further analysis, a statistical criteria can be used when an outlier in standardized deviation terms is greater than 4 or less than –4. The resultant numbers of values omitted are shown below in Table 7.5 by the laboratory type and by the modern standard.

Table 7.5a Number of results omitted by laboratory type

Laboratory type	Number omitted	% of results
AMS	40	19.4
GPC	42	20.3
LSC	124	60.2
All	**206**	**100**

Table 7.5b Number of results omitted by modern standard

Modern standard	Number omitted	% of results
ANU Sucr	20	10.3
Benzene	25	12.9
NBS OXI	52	26.9
NBS OXI/OXII	5	2.6
NBS OXII	66	34.2
Other	25	12.9
All	**193**	**100**

From the tables, it is clear that the majority of results omitted under this criterion are from LSC laboratories and that omission of results is more evenly distributed over the modern standard.

With the removal of the outliers, the distribution of results is more symmetrical.

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

Table 7.6 Descriptive statistics: outliers omitted

Analyses	Number of results	Mean	Median	Min	Max
1	87	–0.227	–0.101	–3.59	3.41
2	210	0.334	0.349	–3.93	3.86
3	103	–0.071	0.100	–3.94	3.61
4	350	–0.1073	–0.1626	–3.8	3.93
Unknown	99	0.116	0.231	–3.86	3.72

A formal analysis of the "laboratory throughput" is shown in Table 7.7 below.

Table 7.7 Effect of number of analyses

```
Source      DF         SS         MS         F         P
Analyses     3      32.38      10.79      4.18     0.006
Error      746    1925.88       2.58
Total      749    1958.26

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  -------+---------+---------+---------
1           87    -0.227     1.610    (---------*--------)
2          210     0.334     1.687                      (------*-----)
3          103    -0.071     1.680         (--------*--------)
4          350    -0.107     1.533            (----*----)
                                    -------+---------+---------+---------
Pooled StDev =     1.607               -0.35      0.00      0.35
```

Table 7.7 shows that there is a statistically significant difference in the average standardized deviation between the different categories of laboratory throughput.

However, we need to also consider that the number of analyses is very strongly related to laboratory type, in that AMS laboratories, in general, tend to have the highest throughput. Therefore, a further analysis, including both laboratory type and throughput, was carried out. The means of the standardized deviations are shown in Table 7.9, cross-classified by both laboratory type and throughput and the formal analysis is summarized in Table 7.8.

Table 7.8 Effect of laboratory type and number of analyses

```
Source      DF      Seq SS      Adj SS      Adj MS      F      P
Technique    2      68.831      69.824      34.912   13.99  0.000
Analyses     3      33.371      33.371      11.124    4.46  0.004
Error      744    1856.056    1856.056       2.495
Total      749    1958.258
```

The formal analysis showed that both the laboratory throughput and laboratory type are significant factors and affect the mean of the standardized deviations as shown in Table 7.9.

Table 7.9 Mean standardized deviation by type and number of analyses

	Nr of analyses				
Laboratory type	1	2	3	4	**All**
AMS	—	–0.7809	–0.3706	–0.3230	**–0.3438**
GPC	1.1584	–0.0591	0.4788	0.5738	**0.2695**
LSC	–0.3293	0.6259	–0.2347	0.5772	**0.2683**
All	–0.2267	0.3338	–0.0711	–0.1073	**0.0073**

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

For each sample and laboratory type, the average standardized deviation can be calculated for all samples. The results are shown in the table and figure below.

Table 7.10 Average standardized deviation for each sample by laboratory type

	AB	C	DF	E	GJ	H	I	**All**
AMS	–0.917	–0.048	–0.355	–0.293	–0.483	0.216	–0.115	**–0.311**
GPC	0.67	–0.131	0.442	0.073	0.301	0.318	0.482	**0.319**
LSC	0.567	0.551	–0.062	0.45	0.182	0.116	–0.034	**0.209**
All	0.065	0.184	–0.077	0.020	–0.078	0.196	0.038	**0.020**

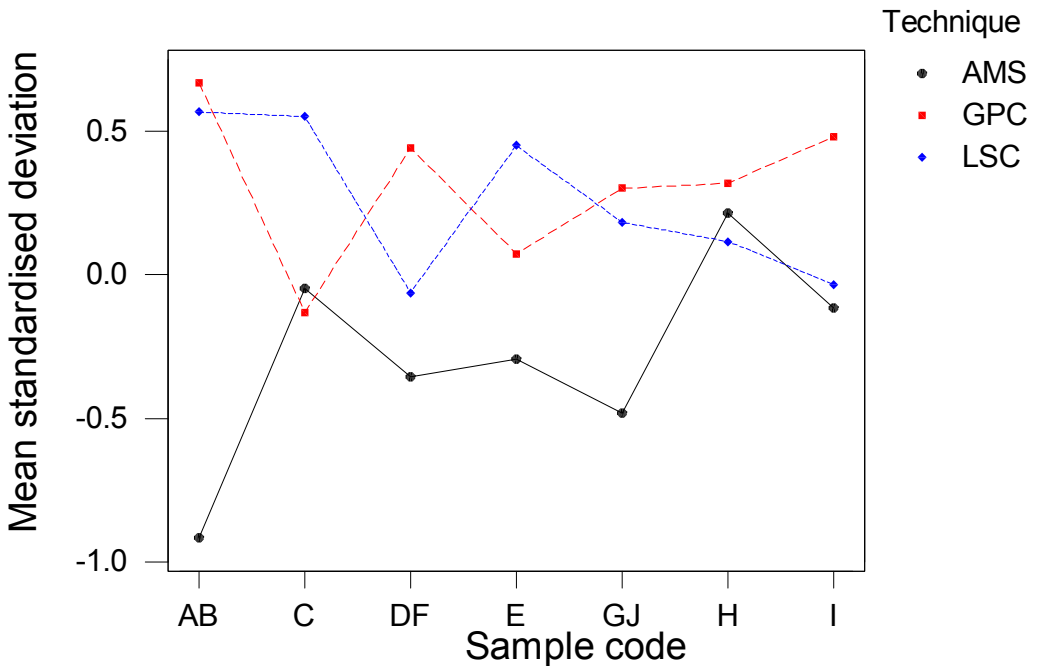

Figure 7.19 Mean standardized deviation by sample and laboratory type

The results for the 3 laboratory types are broadly similar (with the exception of AB, see Section 6) after the omission of outliers and all are generally acceptable (lying in a range of –1 to +1).

7.5 EVALUATION OF LABORATORY ACCURACY

Accepting the consensus values as, in some sense, the true age/activity for each material, we can evaluate the average laboratory difference from the consensus profile. The model used assumes that for a given laboratory there is a potential systematic offset from the consensus profile, which we can estimate, α, see Equation 1. These estimates are summarized in Table 7.11 and shown in Table 7.12.

$$\alpha = (\Sigma[x_i - \mu_i]^2 / s_i^2)\Sigma(1/s_i^2) \quad (1)$$

A summary of the results in Table 7.11. In the 2nd row outliers, offsets >2 or offsets <–2 are excluded.

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

Table 7.11 Summary of offset (pMC) for laboratories

Variable	N	Mean	Median	StDev	Min	Max	Q1	Q3
Offset	92	0.089	–0.010	1.403	–4.5	5.8	–0.3	0.3
Outliers excluded	85	–0.0005	–0.010	0.664	–1.3	1.8	–0.2	0.2

In summary, of the 90 labs for which an uncertainty estimate on the offset could be calculated, 59 were shown to have no offset. The distribution of offsets is shown in Figure 7.20.

Table 7.12 Laboratory offsets in pMC

Lab nr	Number of results	Lower limit on offset	Offset	Upper limit on offset
1	16	–0.24	–0.09	0.05
2	10	–0.19	–0.09	0.02
3	5	–0.71	–0.11	0.48
4	6	–0.77	–0.33	0.11
5	12	1.43	1.79	2.15
6	8	0.12	0.22	0.32
7	8	–0.02	0.29	0.61
8	8	–0.05	0.07	0.20
9	8	–0.21	0.22	0.64
10	6	1.97	2.91	3.86
11	12	–0.75	–0.02	0.72
12	8	–0.66	–0.27	0.12
13	8	–1.56	–0.87	–0.17
14	1	—	–0.50	—
15	11	–0.21	0.01	0.24
16	7	–3.45	–1.16	1.13
17	9	–1.49	–0.91	–0.34
18	8	–1.09	–0.37	0.36
19	8	–0.75	–0.11	0.53
20	9	–0.73	–0.35	0.04
21	8	–0.18	1.55	3.27
22	4	–1.08	–0.16	0.77
23	8	–1.58	–0.91	–0.25
24	7	–0.08	0.21	0.50
25	8	–0.18	0.03	0.23
26	6	1.50	5.81	10.13
27	8	–0.27	–0.11	0.05
28	6	–4.04	–1.27	1.50
29	8	–0.68	–0.40	–0.11
30	8	0.13	0.42	0.70
31	8	–0.21	–0.03	0.15
32	8	–1.17	–0.51	0.15
33	8	0.00	0.47	0.94
34	8	0.07	0.11	0.14
35	7	–1.76	–1.33	–0.91
36	10	–0.09	–0.01	0.07
37	13	–0.17	–0.07	0.04

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

Table 7.12 Laboratory offsets in pMC *(Continued)*

Lab nr	Number of results	Lower limit on offset	Offset	Upper limit on offset
38	9	0.02	0.12	0.22
39	8	–2.52	–1.33	–0.14
40	7	0.02	0.16	0.29
41	10	–0.08	0.07	0.21
42	6	0.07	1.69	3.31
43	8	–1.91	–1.06	–0.20
44	8	0.38	1.82	3.25
45	3	–3.14	–0.16	2.82
46	8	–0.28	–0.11	0.06
47	8	–0.65	–0.32	0.00
48	8	–0.20	0.07	0.35
49	22	–0.17	–0.07	0.03
50	16	–0.09	–0.02	0.05
51	28	–0.03	0.11	0.25
52	8	–0.28	–0.06	0.15
53	8	–11.05	–4.45	2.15
54	8	–0.05	0.30	0.65
55	8	–0.37	–0.12	0.12
56	6	–3.63	–2.56	–1.50
57	7	0.10	0.69	1.28
58	8	–0.40	–0.15	0.11
59	8	–2.26	–0.67	0.92
60	8	–0.29	–0.09	0.12
61	7	–0.55	–0.23	0.09
62	6	0.04	0.38	0.72
63	2	–8.15	–1.01	6.13
64	8	–0.09	0.08	0.26
65	8	0.05	0.13	0.22
66	8	0.30	0.48	0.65
67	5	–2.99	–1.05	0.88
68	8	–1.82	–0.94	–0.05
69	7	–7.31	–4.14	–0.97
70	16	1.49	4.85	8.22
71	7	0.12	0.97	1.81
72	10	–0.17	0.15	0.47
73	8	–0.15	0.01	0.17
74	11	–0.01	0.13	0.27
75	7	–0.25	0.36	0.98
76	8	–0.06	0.46	0.97
77	10	0.02	0.22	0.42
78	7	–3.68	0.10	3.89
79	8	–0.17	–0.05	0.06
80	1	—	5.80	—
81	8	–3.24	–1.01	1.22
82	8	–0.54	–0.28	–0.01

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

Table 7.12 Laboratory offsets in pMC *(Continued)*

Lab nr	Number of results	Lower limit on offset	Offset	Upper limit on offset
83	8	–0.10	0.01	0.12
84	20	–0.24	–0.09	0.06
85	8	0.34	0.43	0.51
86	8	–0.07	0.22	0.52
87	8	0.16	0.40	0.65
88	18	–0.17	–0.01	0.15
89	8	–0.53	0.45	1.42
90	12	0.94	1.55	2.17
91	8	–0.09	0.05	0.18
92	3	1.07	1.77	2.46

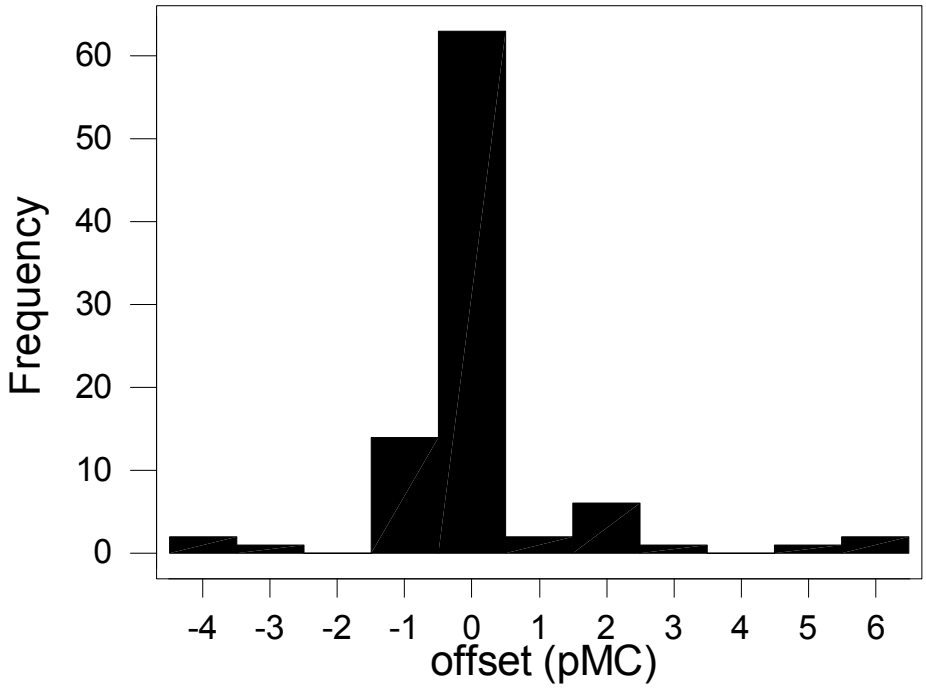

Figure 7.20 Distribution of laboratory offset relative to consensus values

Another possible calculation of offset can be based on the dendro-dated samples, of which 3 were included specifically for this purpose. Assuming a known age for these samples (based on the master chronology), an offset for each laboratory can then be estimated.

7.5.1 Offset Relative to the Dendro-Dated Wood Samples (yr BP)

A total of 4 dendro-dated wood samples were included in the list of core samples. They were Samples D and F (duplicates) from the Belfast master chronology, dendro-dated to 3200–3239 BC. Sample I (also from the Belfast master chronology), dendro-dated to 3299–3257 BC. Sample H was from the German oak chronology and was dendro-dated to 313–294 BC. A simple, exploratory summary of the findings and their comparison with the master calibration results is described in the following.

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

FIRI Samples D and F

Dendro-dated to 3239–3200 BC, this sample is linked to 4 samples on the master chronology. The average of the ^{14}C ages gives a "true" age of 4495 BP.

Table 7.13 Linked master calibration samples

Decadal midpoint	^{14}C age (1 σ)
3205	4528 ± 18
3215	4497 ± 11
3225	4495 ± 18
3235	4461 ± 18

FIRI Sample I

Dendro-dated to 3299–3257 BC, this sample is linked to 5 samples on the master chronology.

Table 7.14 Linked master calibration samples

Decadal midpoint	^{14}C age (1 σ)
3255	4455 ± 18
3265	4486 ± 18
3275	4480 ± 18
3285	4469 ± 18
3295	4468 ± 18

An average of the ^{14}C ages gives a "true" age of 4471 BP.

FIRI Sample H

Dendro-dated to 313–294 BC, this sample links to 3 samples on the master chronology.

Table 7.15 Linked master calibration samples

Decadal midpoint	^{14}C age (1 σ)
315	2210 ± 25
305	2211 ± 25
295	2225 ± 18

An average of the ^{14}C ages gives a "true" age of 2215 BP.

Similarly, using the master chronology ^{14}C ages as the "true" age for each laboratory, it is possible to estimate the systematic offset (if any) relative to these "true ages." However, it should be pointed out that, in fact, the consensus values for these samples are only slightly different from those extracted from the master calibration curve (4495 versus 4508 yr BP for DF, 4471 versus 4485 yr BP for I, and 2215 versus 2232 yr BP for Sample H).

Summarizing the offsets, we have:

Table 7.16 Offset in yr BP from the master ^{14}C ages

	N	Mean	Median	Min	Max	Q1	Q3	StDev
Offset	90	16.8	17.0	–642	414	–22	74	140.0
Outliers excluded	81	27	17	–218	209	–17	72	81

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

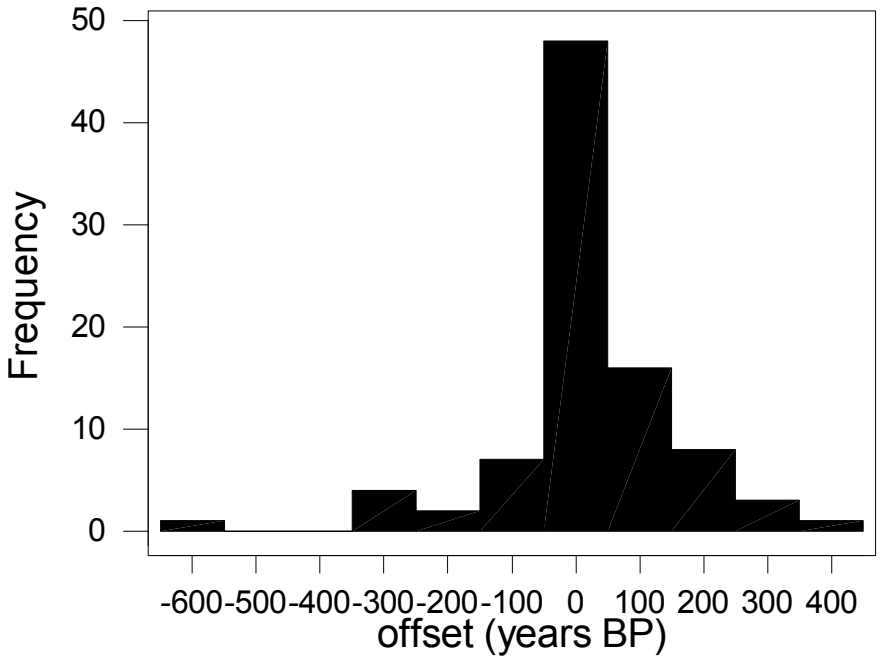

Figure 7.21 Distribution of offsets relative to the dendro-dated samples

Table 7.17 Lab offset (yr BP), based only on the dendro-dated samples (DF, I, H)

Lab nr	Number of results	Lower limit on offset	Offset	Upper limit on offset
1	8	48.83	80.86	112.89
2	6	6.53	37.24	67.96
3	2	–150.51	137.00	424.51
4	4	–31.83	89.87	211.56
5	7	–306.67	–218.46	–130.24
6	4	–46.34	–0.97	44.39
7	4	–135.38	–66.25	2.88
8	4	–46.65	–6.50	33.65
9	4	12.43	70.53	128.63
10	4	–420.29	–345.04	–269.79
11	7	–211.17	–62.76	85.65
12	4	–67.22	43.50	154.22
13	3	–84.91	199.67	484.24
15	7	–43.11	–6.04	31.03
16	3	94.17	203.95	313.73
17	5	–15.97	18.35	52.67
18	4	–140.97	12.64	166.25
19	4	1.11	62.13	123.14
20	5	73.72	109.46	145.21
21	4	–793.37	–327.44	138.50
22	3	–323.13	16.58	356.28

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

Table 7.17 Lab offset (yr BP), based only on the dendro-dated samples (DF, I, H) *(Continued)*

Lab nr	Number of results	Lower limit on offset	Offset	Upper limit on offset
23	4	75.46	209.49	343.52
24	4	–74.40	–14.28	45.84
25	4	–7.91	56.00	119.91
26	4	–1144.10	–642.19	–140.28
27	4	10.07	37.73	65.39
28	3	–626.91	–63.44	500.02
29	4	–39.33	75.73	190.79
30	4	–174.60	1.00	176.60
31	4	–70.59	8.44	87.46
32	4	–77.49	72.58	222.66
33	4	–147.00	–60.64	25.72
34	4	–94.50	–33.88	26.75
35	4	–62.96	104.21	271.38
36	6	–7.38	12.68	32.74
37	6	6.09	35.28	64.47
38	5	–41.22	–18.21	4.80
39	4	209.19	258.46	307.74
40	4	–51.43	–7.69	36.05
41	6	–20.89	10.32	41.54
42	4	–603.20	–321.25	–39.31
43	4	64.38	205.79	347.20
44	4	–257.58	–129.90	–2.21
45	2	–97.79	–35.86	26.06
46	4	–13.35	36.63	86.62
47	4	–31.27	125.28	281.82
48	4	–67.40	34.49	136.37
49	11	–11.40	17.41	46.23
50	8	–9.92	21.45	52.81
51	16	3.39	31.97	60.54
52	4	35.24	77.15	119.06
53	4	–1732.51	56.91	1846.32
54	3	–94.28	–20.99	52.30
55	4	–68.55	4.44	77.44
56	4	56.51	262.05	467.58
57	4	–226.99	–106.32	14.35
58	4	–29.62	42.03	113.68
59	4	–215.48	276.36	768.20
60	4	–36.92	42.28	121.49
61	4	–51.10	21.36	93.81
62	1	—	29.00	—
63	2	–873.23	160.00	1193.23
64	4	47.51	73.50	99.49
65	4	–39.40	–16.57	6.27
66	4	–50.75	–27.75	–4.75
67	4	–255.02	163.26	581.53

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press

Table 7.17 Lab offset (yr BP), based only on the dendro-dated samples (DF, I, H) *(Continued)*

Lab nr	Number of results	Lower limit on offset	Offset	Upper limit on offset
68	4	169.60	192.55	215.50
69	4	26.45	414.43	802.42
70	8	–143.13	118.80	380.73
71	4	–245.85	–105.83	34.19
72	5	–87.53	20.01	127.55
73	4	–31.99	–1.17	29.65
74	7	–5.51	15.91	37.33
75	4	–189.71	–15.24	159.22
76	4	–36.24	–1.50	33.24
77	7	–60.72	–30.97	–1.22
78	4	–733.60	–27.86	677.87
79	4	–2.26	31.22	64.70
81	4	–116.91	182.93	482.76
82	4	28.77	102.39	176.01
83	4	–27.16	0.37	27.90
84	8	–37.68	–0.83	36.02
85	4	–55.00	–35.27	–15.53
86	4	–131.05	–26.97	77.12
87	4	–100.49	–49.26	1.98
88	9	–49.90	–16.60	16.70
89	4	–28.03	44.21	116.45
90	8	–239.84	–153.61	–67.37
91	4	–48.51	–7.57	33.36
92	1	—	–251.00	—

7.6 CONCLUSIONS

Consensus values (and their error) for the FIRI samples have been derived. Concerns remain over the consensus value for the Kauri wood sample due to the reporting difficulties for this sample. The sensitivity of the results to different calculation algorithms has been shown to be small (with the exception of the Kauri wood). Consensus values for the dendro-dated wood samples are very close to the values derived from the master calibration curve, which adds confidence in the results derived.

When considering laboratory performance, we have evaluated standardized deviations from the consensus values and have shown that these can be linked to the laboratory type. Calculation of the offsets has also shown that more than half the laboratories have no systematic offset, and that those laboratories that have a systematic offset, generally have small offsets (with only a few exceptions). Laboratories received this information for their consideration and, thus, were able to explore any causes, and then instigate any necessary corrective actions.

https://doi.org/10.1017/S003382220003263X Published online by Cambridge University Press