Scott, E.M., Cook, G.T., Naysmith, P. and Staff, R.A. (2019) Learning from the wood samples in ICS, TIRI, FIRI, VIRI and SIRI. *Radiocarbon*, (doi:10.1017/RDC.2019.12)

Learning from the wood samples in ICS, TIRI, FIRI, VIRI and SIRI.

*E M Scott, G T Cook, P Naysmith, R Staff

*School of Mathematics and Statistics, University of Glasgow, Glasgow

Scottish Universities Environmental Research Centre, University of Glasgow, Glasgow

Corresponding author: marian.scott@glasgow.ac.uk

Abstract: Each of the laboratory inter-comparisons (from ICS onwards) has included wood samples, many of them dendrochronologically dated. In the early years, as a result of the majority of laboratories being radiometric, these samples were typically blocks of 20-40 rings, but more recently (SIRI), they have been single ring samples. The sample ages have spanned background through to modern. In some inter-comparisons, we have examined different wood pre-treatment effects, in others the focus has been on background samples. In this paper, we illustrate what we have learnt from these extensive inter-comparisons involving wood samples and how the results contribute to the global IntCal effort.

Keywords:

Introduction

In the 1980's, with the growth in the number of radiocarbon laboratories, including new AMS laboratories, a proposal was made for a formal quality assurance program to be introduced (Long and Kalin (1990)). This could take the form of a laboratory inter-comparison or proficiency trial as set out in Thomson et al (2006), where a selection of samples is chosen to be used in the inter-comparison and all working laboratories are invited to take part in the inter-comparison to check their own individual performance. Following from early work, a community programme of inter-comparisons began (Scott et al, 2018). The samples selected to be used in these programmes were natural and routinely dated materials, many of which had the potential to become internationally recognised reference materials. The main criteria for selecting samples were that they should:1) Be of archaeological and/or geological interest, 2) Cover the broad spectrum of laboratory experience (age, sample type, etc), 3) Satisfy rigorous homogeneity testing, 4) Be known age if possible. In this short paper we concentrate on the wood samples relevant to criteria 1,2 and 4, used in the inter-comparison studies. We will briefly describe the pre-treatment method used to extract holo cellulose, and the connections between the different inter-comparisons where the same material has been used on several occasions (as wood or cellulose). Where appropriate, updated consensus values will be provided. Finally, we provide an illustration of the benefits which an individual laboratory can gain from a well characterised inter-comparison sample.

2. Samples and studies

2.1The different wood samples and their pretreatment to holo cellulose

We now reflect on the compendium of wood samples that have been used in the inter-comparisons starting from ICS in 1988. Table 1 describes the 29 wood samples including cellulose, that have been used. This paper will not consider any further the near background or background wood samples, namely FIRI A and B (Kauri), VIRI K (Hohenheim), SIRI A and L (Hohenheim and Oregon) or TIRI G (close to background). Similarly there will be no further discussion of modern samples, VIRI O (FIRI K) and IAEA cellulose (TIRI C). SIRI M while used previously will not be further considered since in SIRI this was provided only to radiometric labs.

Fifteen of the wood samples have been dendro-dated. Dendro-dated woods are valuable to include not least since they provide an independent measure of the age of the sample (known calendar age). This provides an opportunity to compare the results with the known age (after calibration) (and so is a more nuanced comparison than simply using the consensus ^{14}C age), and allows laboratories to directly connect to some of the ongoing calibration work. Several of the samples had also been previously dated. Historically, to ensure that we had sufficient materials, the samples have been provided as blocks of rings (either 20 or 40 rings), and we have chosen the blocks to lie on a ‘plateau’ on the calibration curve. As a result, there has been no formal homogeneity testing. In SIRI where the focus was on AMS laboratories, we have provided for the first time single rings.

Study	**Sample Code**	**Sample Type**	**Pre-treatment prior to shipment to labs**	**Consensus (^{14}C BP)**
ICS	Stage 2	Belfast pine (241-260BC) (in duplicate)	cellulose extraction	2278±32*
ICS	Stage 3	Belfast pine (221-240BC) (in duplicate)	None	2208±13*
	Stage 3	1521-1550AD	None	297±22*
	Stage 3	1841-1870AD	None	109±21*
TIRI	Sample B	Belfast pine (3239-3200BC)	None	4503 ± 6
	Sample J	Crannog (optional)	None	1605 ± 8
	Sample C	IAEA cellulose (modern)	None	**-**
	Sample G	Fuglaness (close to background)	None	**-**
FIRI	Sample D	Belfast wood (pine) 3239-3200BC, **TIRI B**	None	4508 ± 3

	Sample F	Belfast wood (duplicate of D)	None	4508 ± 3
	Sample H	German wood (oak 313-294BC)	None	2232 ± 5
	Sample I	Belfast cellulose (3299-3257BC)	Cellulose extraction	4485 ± 5
	Sample L	Dogee Barrow (optional)		2505±39
	Sample K	Cambridge cellulose(1820-1880AD) Optional sample	Cellulose extraction	**-**
	Sample A	Kauri background		**-**
	Sample B	Kauri background (duplicate to A)		**-**
VIRI	Sample L	Wood (Corlea Q5994) 221-260BC	None	2234 ± 17
	Sample O	Cellulose (1820-1880AD) **FIRI K**	Cellulose extraction	125 ± 16
	Sample M	Loch Tay crannog (oak)		2430 ± 16
	Sample N	Loch Tay crannog (alder)		2437 ± 17
	Sample K	Hohenheim background		**-**
SIRI	Sample F	Wood (Belfast) 1487AD	None	363 ± 3
	Sample G	Wood (Belfast) 1479AD	None	377 ± 5
	Sample H	Wood (Belfast) 1475AD	None	386 ± 3
	Sample E	Kauri YD	None	10843 ± 6
	Sample I	Lake Gribben YD	None	9995 ± 5
	Sample A	Hohenheim background	None	**-**
	Sample L	Oregon background	None	**-**
	Sample M	Crannog wood	none	**-**

- [*] The ICS results are summarised as the overall mean and standard error of the mean

Table 1: summary values for all wood and cellulose samples.

2.2 Wood descriptions

In ICS, the samples were provided by Professor M Baillie, Queens University, Belfast, comprising two samples of contiguous 20 rings of dendro-dated bog oak. TIRI B was Scots pine (*Pinus sylvestris*) collected by Professor Ballie in December 1991. It grew on the western side of the Gary Bog, County Antrim and was designated Q7780. Each sample was a block of 40 rings, representing growth rings 74-113 of the 347-year tree. The sample conforms exactly to two of the bidecadal samples of oak used in the original high precision calibration (Pearson & Stuiver 1986). This sample was dendro-dated from 3200BC to 3239BC. The TIRI J timber was in the form of a large morticed baulk, lying just behind the outer palisade of Buiston Crannog near Kilmaurs, Ayrshire (NGR 4154 4351). Although no longer *in situ*, it resembled the mortice planks used to secure the stakes of the outer palisade and is interpreted here as having formed part of the latter. The sample was supplied by Dr B A Crone of AOC Archaeology.

FIRI D and F were identical to TIRI B. FIRI I was a second bulk Scots Pine sample from Garry Bog, supplied by Professor M Ballie. He supplied 16.3kg of Scots Pine which had a finite 40-yr ring span, and again had the sample identification number Q7780. The dendro-dated age span was 3299-3257 BC. The FIRI H sample was provided by the Dr M Spurk of the University of Hohenheim comprising 9.6kg of dendro-dated oak. The sample identification number was Pettstadt 262. The sample had 20 annual growth rings dating from 313 BC to 294 BC. FIRI L was a wood sample (part of a log) of approximately 10kg covering annual rings from the burial mound of Dogee Barrow, grave 8, (the Tuva king barrows from Ssythia) was provided by Dr G Zaitseva of the Institute of the History of Material Culture. The material was excavated in 1998 and was very degraded. Its approximate age was 2300-2400 BP. FIRI K was oak (*Quercus robur*), obtained from Dr R Switsur of the Goodwin Institute for Quaternary Research. The tree was planted around AD 1722 and the material corresponding to the period AD 1820-1880 (a relatively flat area on the calibration curve) was removed to provide a sample of 10.4kg.

VIRI L was again provided by Professor M Ballie. This sample is identified as Corlea Q5994. Samples M and N were provided by Professor G Cook, SUERC. Sample M is an oak sample and sample N is an alder and they come from a crannog site at Loch Tay, Scotland.

The SIRI samples F, G, and H were single ring samples again provided from the Queens University of Belfast. SIRI E is kauri and was provided by Professor A Hogg, Waikato University, New Zealand. It is a decadal sample and its code is Tawa YD Kauri wood rings 1251-60. SIRI I was provided by Professor I Panyushkina of the University of Arizona.

2.2 Wood and Cellulose pre-treatment

Whole Wood: Many of the samples came from dendrochronology laboratories and were simply cut into suitable sized fragments for distribution. For others, the samples were digested in 0.5M KOH at 80^{0}C, soaked in distilled water to remove excess alkali and then digested in hot 2M HCl. Finally, the wood was again soaked in distilled water to remove excess acid and dried to a constant weight in a vacuum oven.

Holo-Cellulose: The wood was either chopped into small pieces, or shavings were produced using a power plane. The material was then subjected to repeated digestion in 2M potassium hydroxide, washing, acidification and bleaching in sodium chlorite/hydrochloric acid solution. The fibrous extract was washed free of chlorite with distilled water, oven dried at 40^{0}C and thoroughly mixed by tumbling

2.3 The intercomparison studies

A brief summary of the studies where wood and cellulose are used is given below (full details can be found in Scott et al, 2018).

ICS (Cook et al, 1990, Harkness et al, 1989, Scott et al, 1989, 1990, 1991): In this three stage trial, one of the goals was the quantitative assessment of variability at different stages in the dating process. In Stage 2 we provided a cellulose sample (in duplicate) and in Stage 3, the 3 wood samples were provided (one in duplicate). All three samples had associated dendro-dates, and one was the contiguous 20 rings to the cellulose sample in stage 2. Following the ICS study, **TIRI** (the Third International Radiocarbon Inter-comparison) (Scott et al., 1992, Scott 2003) included one dendro-dated sample in addition to the IAEA cellulose (C4) and two other wood samples, one >30K. The next study in the sequence was **FIRI** (the Fourth International Radiocarbon Inter-comparison) which was completed in 2000. FIRI included an extensive set of wood samples (including background samples), and one sample that had been used in TIRI. The Fifth International Radiocarbon Inter-comparison (**VIRI)** commenced in 2004 and included cellulose, dendro-dated wood, background wood and several other wood samples. VIRI L spanned the 40 rings comprising ICS2 and ICS3 (Scott et al, 2010) The most recently completed exercise is **SIRI** (the Sixth International Radiocarbon Inter-comparison), which commenced in 2013 and was completed in 2016, including 8 different wood samples, 3 of which were single dendro-dated rings from a 30-year sequence (Scott et al, 2017).

2.3 Statistical analysis

Our approach has been first to assess the distribution of results, identifying any outliers, before proceeding to evaluate laboratory performance (in terms of bias and error multipliers both internal and external (Aitchison et al, 1990) and to quantify the consensus value for each material (including uncertainty) (Scott et al, 2018). In this paper, we also consider the Chi-squared statistic to evaluate uncertainty relative to that expected given the quoted errors.

3. Results and Discussion

Table 1 presents the combined reference information for all wood samples including their codes and published consensus values. For those samples used in ICS and as optional in TIRI and FIRI, we have simply reported here the mean and standard error since typically there were insufficient numbers of results to confirm a consensus value. Figure 1 shows the boxplot of the distribution of results for all 20 wood and cellulose samples, spanning modern to 5,000 BP approximately, excluding SIRI E and SIRI I at 10,000BP.

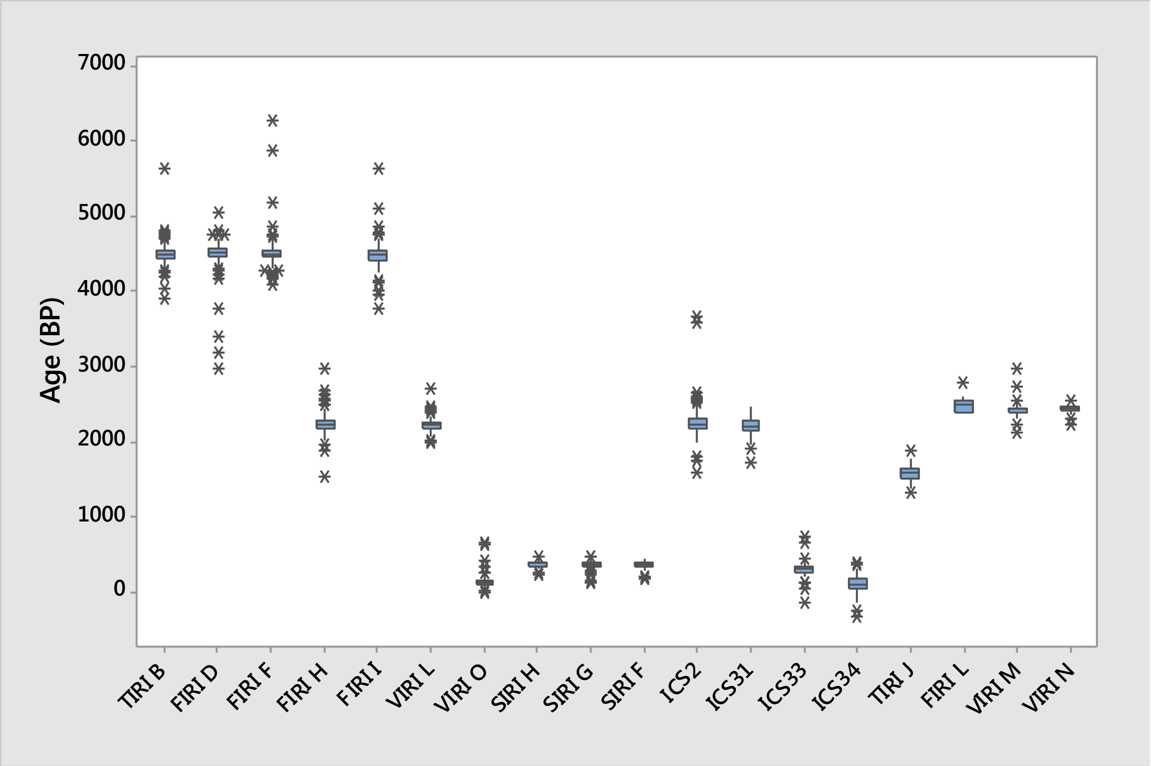

Figure 1: Boxplot of the distribution of results for all wood and cellulose samples.

It is natural to consider as well as the reported ages, the uncertainty associated with each result and table 2 shows basic summaries for the quoted errors. We can see a clear difference in the magnitude of the quoted errors from VIRI onwards, with generally decreasing uncertainties, (moderated of course by the age of the sample), though the minimum uncertainties remain unchanged.

Table 2: summary of quoted errors

sample	Number of observations	Median error (yrs)	Minimum (yrs)	Maximum (yrs)
ICS21	39	50	19	260
ICS22	39	55	20	430
ICS31	43	50	18	230
ICS32	43	50	18	120
ICS33	41	50	13	140
ICS34	43	50	13	150
TIRI B	79	60	17	190
TIRI J	36	45	10	82
FIRI D	108	43	10	240
FIRI F	103	50	16	290
FIRI H	97	40	19	220
FIRI I	94	50	20	290
FIRI L	10	48	25	202
VIRI L	57	30	12	71

VIRI O	63	30	13	148
VIRI M	55	30	12	104
VIRI N	34	30	17	50
SIRI H	73	24	15	64
SIRI G	79	24	14	63
SIRI F	79	24	15	63
SIRI E	73	42	20	93
SIRI L	76	40	20	75

3.1 Specific comparisons and investigations

In this section, we focus on the dendro-dated samples, broken into the distinct time periods, and specifically using the linked samples, including ICS2 and ICS3 which cover the same span of rings as VIRI L, and TIRI B and FIRI D and F which are the same sample, and SIRI F,G and H which are single rings from a span of 32 years. We also consider the results in the context of IntCal13, evidencing the variability that is apparent across laboratories measuring the same material (both decadal blocks as well as single rings). We show three examples

3.1.1 FIRI H, VIRI L, M and N, ICS2 and ICS3 (Period 350-220BC)

Table 3 shows the basic summaries for the 5 samples in this period, noting that VIRI L is a 40 ring block spanning the contiguous 20 rings blocks for ICS2 and ICS3. As expected, the mean VIRI L age lies within the ICS2 and ICS3 age range, VIRI L results are completely consistent with the ICS3 results in terms of scatter. Figure 2 shows FIRI H, VIRI L, M and N plotted on IntCal13 (Reimer et al, 2013).

sample	N	Mean (^{14}C BP)	SE mean	Std deviation	Minimum	Maximum
FIRI H	98	2246.4	16.3	161.7	1530.0	2980.0
VIRI M	55	2437.2	15.3	113.5	2120.0	2990.0
VIRI N	34	2433.2	9.36	54.6	2237.0	2540.0
VIRI L	57	2225.3	14.9	112.2	1990.0	2702.0
ICS2	78	2278.2	31.8	281.1	1600.0	3680.0
ICS2*	76	2242.6	20.2	175.9	1600.0	2670.0
ICS3	86	2208.5	12.5	115.6	1740.0	2460.0

Table 3: Summary for samples in period 350-220BC (* two values removed as outliers, one pair of duplicates)

From Figure 2, the results are distributed well around the calibration curve, but we can see the wide dispersal of dates beyond the curve uncertainty (1 sigma) band but with the bulk of the measurements lying within the band. (Note ICS2 and 3 are not plotted on the curve).

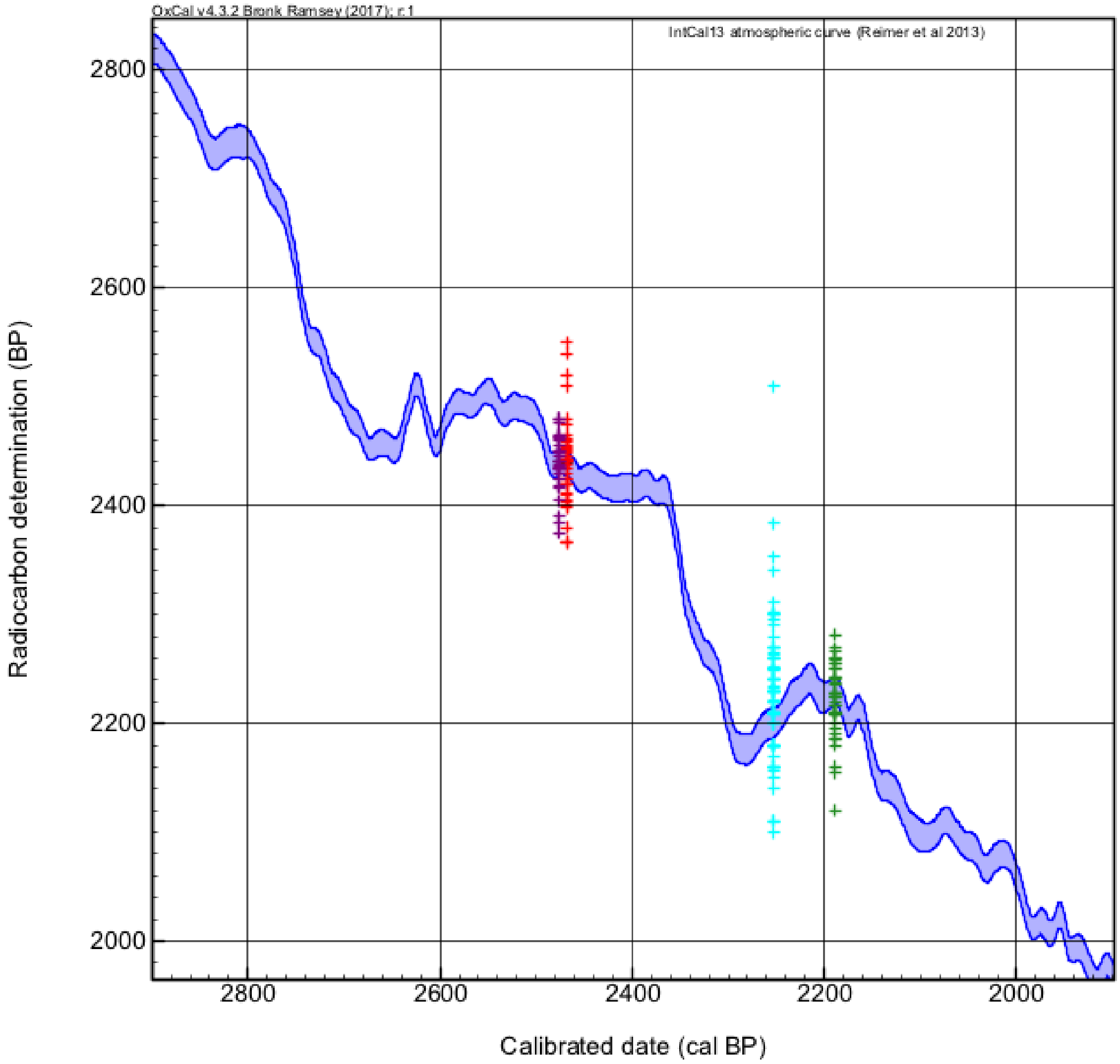

Figure 2: IntCal13 (1 standard deviation envelope) and intercomparison sample scatter.

3.1.2 Period 3239-3200BC

sample	N	Mean (^{14}C BP)	SE mean	Std deviation	minimum	maximum
TIRI B	80	4502.5	21.3	190.2	3900.0	5640.0
FIRI D	108	4487.3	24.9	258.6	2990.0	5060.0
FIRI D*	105	4524.3	13.2	135.4	3790.0	5060.0
FIRI F	103	4537.0	25.6	259.3	4100.0	6270.0
FIRI F*	101	4506.6	14.0	140.9	4100.0	5178.0
FIRI I	94	4494.8	21.5	208.1	3780.0	5650.0

Table 4: Summary for samples in period 3239-3200BC (* rows represent results for FIRI D and FIRI F after removal of 2 outliers)

Table 4 shows the basic summaries for the 4 samples in this period, noting that TIRI B and FIRI D and F are identical 40 ring blocks. Figure 3 shows the results plotted on IntCal13 (note TIRI B not shown). There is considerable variability evident in the FIRI results, however with the removal of 2 outliers in each set, the variation (standard deviations) for FIRI D and FIRI F are lower than for TIRI B.

From Figure 3, similar features are observed, with results distributed well around the calibration curve, but we can see the wide dispersal of dates beyond the curve uncertainty (1 sigma) band.

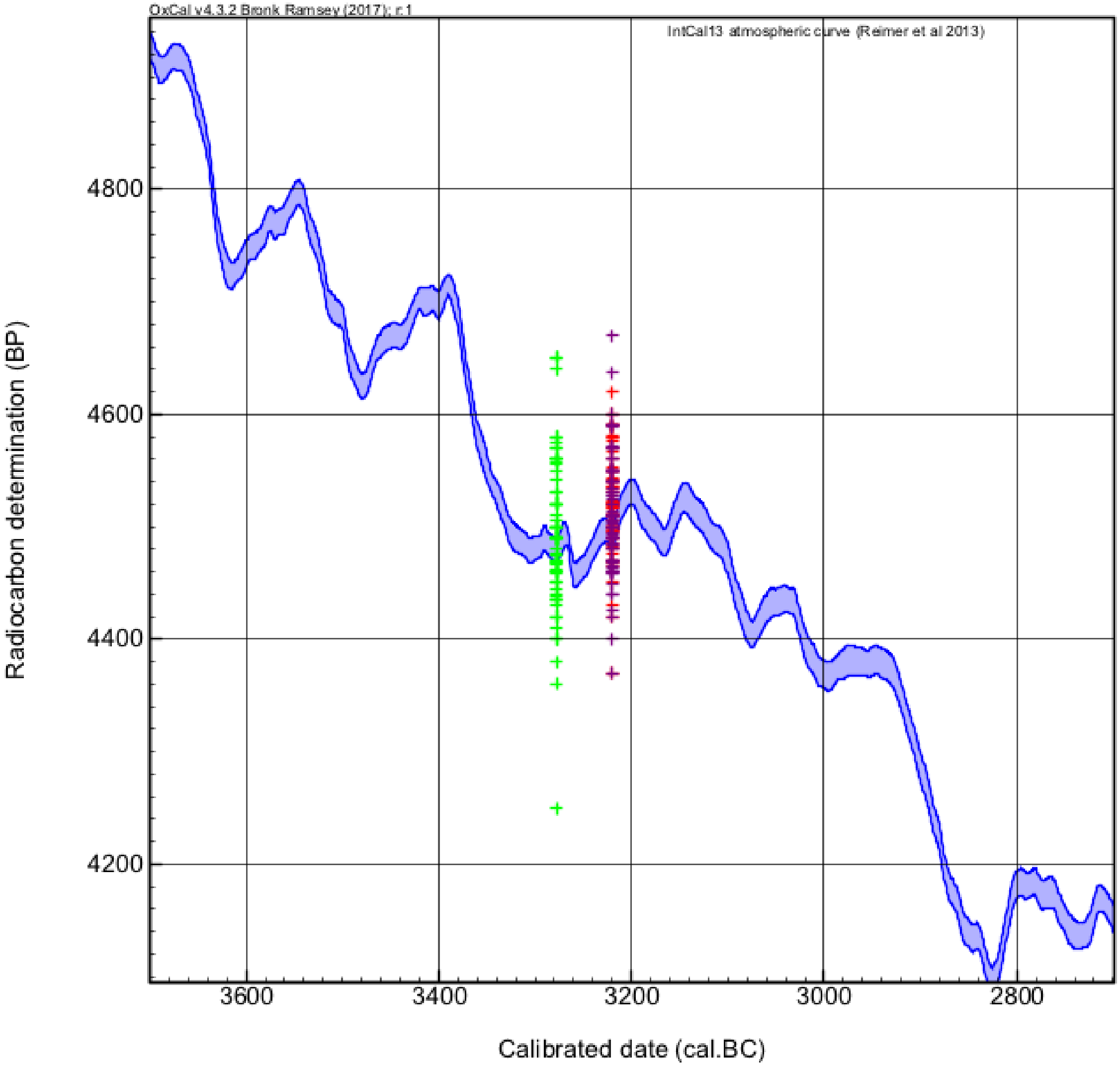

Figure 3: IntCal13 curve and (1 standard deviation envelope) and intercomparison sample scatter

3.1.3 wood 1475-1180AD, samples VIRI O, SIRI F,G and H

sample	N	Mean (^{14}C BP)	Se Mean	Std deviation	minimum	maximum
VIRI O	63	148.4	14.4	114.1	10.0	667.0
SIRI H	74	381.53	4.83	41.51	245.00	486.00
SIRI G	80	367.52	6.96	62.29	128.00	478.00
SIRI F	80	361.65	5.53	49.43	184.00	461.00

Table 5: Summary for samples in period 1475-1180AD

Table 5 shows the basic summaries for the 4 samples in this period, noting that the SIRI samples are single rings while VIRI O spans 60 rings. Figure 4 shows VIRI O, SIRI F, G and H plotted on IntCal13. From Figure 4, similar features are observed as in Figures 2 and 3 , with results distributed well around the calibration curve, but we can see the wide dispersal of dates beyond the curve uncertainty band. While some of this scatter must be due to the spread in age (20 or 40 rings blocks), there is evidence of a reduction in this scatter when we consider the 3 single ring samples, the standard deviation of results is reduced by approximately 2.

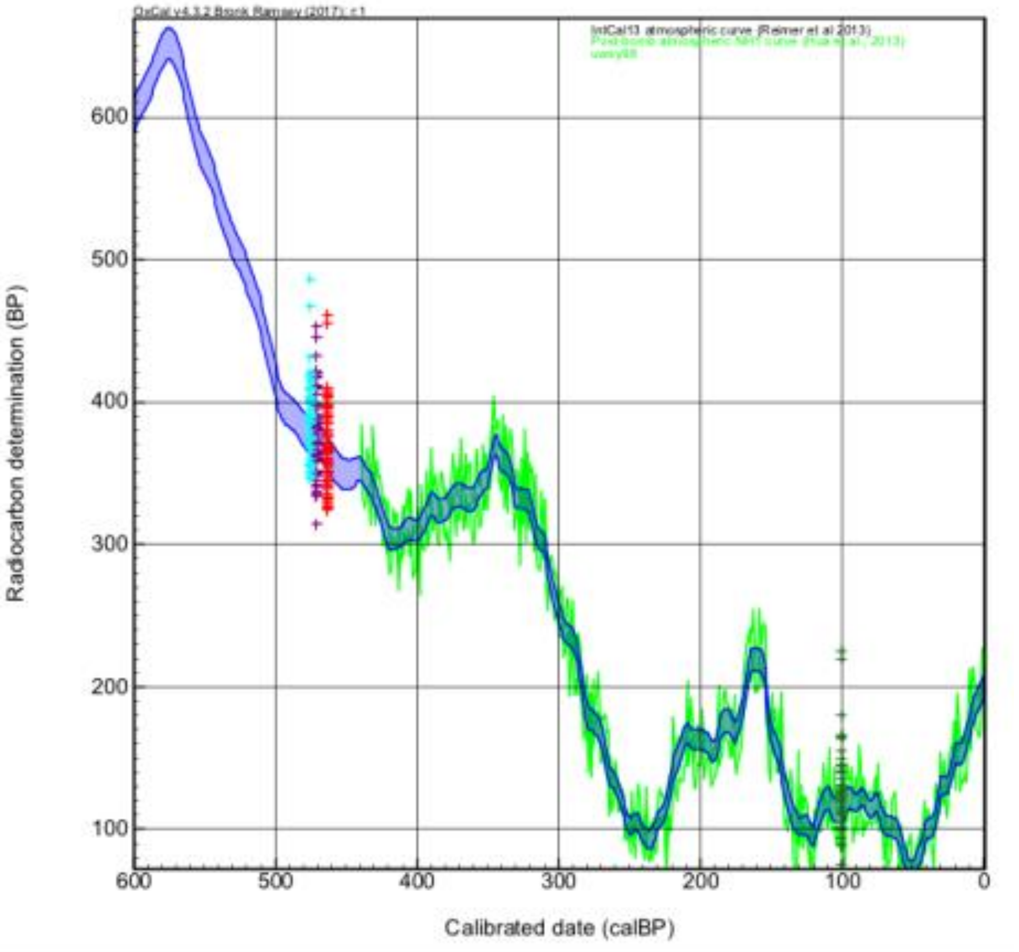

Figure 4: IntCal13 (1 standard deviation envelope), single year ^{14}C data set (Stuiver et al, 1998) and SIRI sample scatter, where the green curve is the University of Washington 1998 single year 14C dataset (Stuiver, M., Reimer, P.J., and Braziunas, T.F., (1998))

3.2 Excess variation

Traditionally, evaluation of z-scores, is a standard approach to evaluate the performance relative to the consensus value estimated using the same procedure as described in(Scott et al, 2018), but of particular interest in this context is the variability in the results and checking of the measurement uncertainties, so in this context we use a zeta score and evaluate the reduced Chi-squared statistic. The zeta score is defined below and is interpreted similar to the z-score where. x_m, the reported result, x_A, the assigned or true value for the material, σ_p, the target value for standard deviation and in addition σ_a is the uncertainty on the consensus value.

$$\text{zeta-score}= (x_m-x_A)/ \sqrt{(\sigma_p^2+ \sigma_a^2)}$$

The target value for the standard deviation is sometimes called the ‘standard deviation for proficiency testing’ is sometimes taken as the standard uncertainty that is regarded as optimal for the application purpose. (Analytical Methods Committee, AMCTB No. 74, 2016)

Interpretation of the zeta-scores is similar to z-scores as

- $|\text{zeta-score}| \leq 2$ result is considered satisfactory
- $2 < |\text{zeta-score}| < 3$ warning, evaluate the result
- $|\text{zeta-score}| \geq 3$ action, this result is anomalous

It is also common to evaluate a reduced χ^2 (sometimes also called the MWSD). The reduced χ^2 is the χ^2 divided by n-1 (where n is the number of observations used in the calculation of the consensus value). We compare the reduced χ^2 value to 1, values greater than 1 would indicate over dispersion in the results around the consensus value. Figure 5 shows the zeta-scores (which include the uncertainty on the consensus value) in a probability plot (to check linearity) which shows some evidence of measurement uncertainty being under-estimated (given the deviations from linearity in the tails).

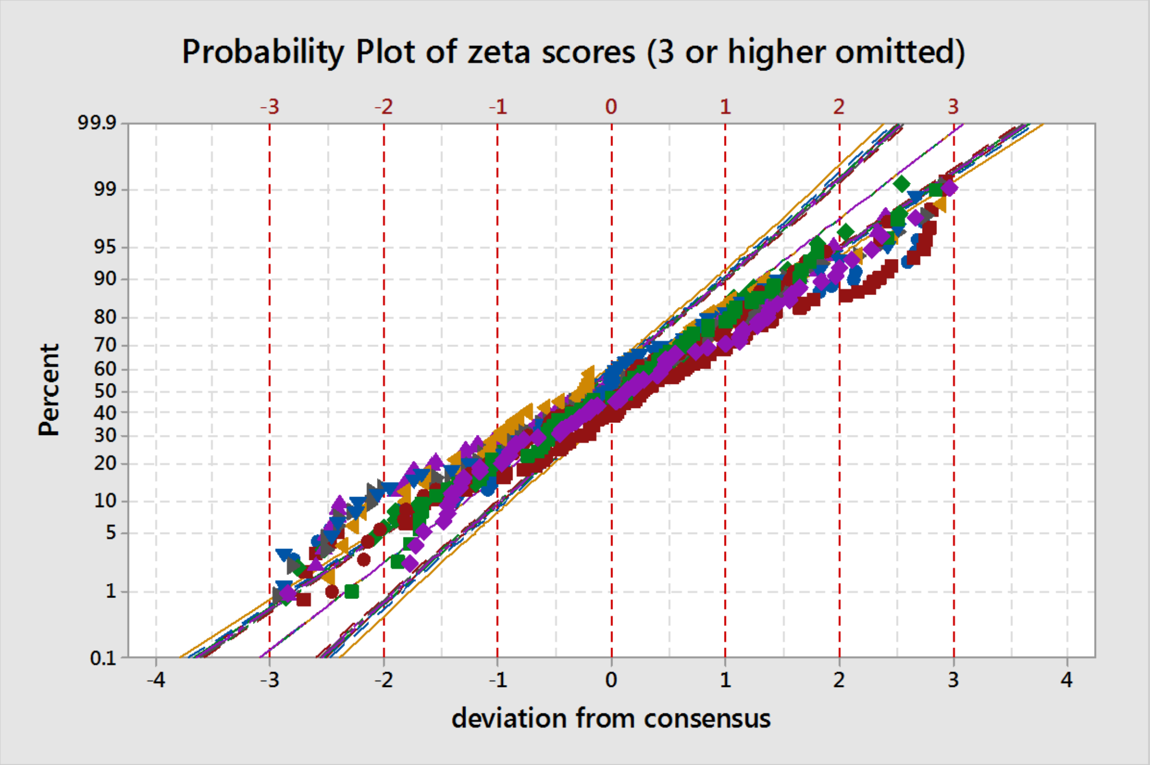

Figure 5: Probability plot of deviations from consensus (each colour represents a different sample)

sample	N	Chi-squared	Reduced Chi-squared
TIRI B	68	113.64	1.67
FIRI D	92	177.43	1.93
FIRI F	82	115.12	1.40
FIRI H	80	151.71	1.89
FIRI I	81	151.37	1.87
VIRI L	46	76.76	1.67
VIRI O	58	101.11	1.74
SIRI H	68	97.25	1.43
SIRI G	69	87.86	1.27
SIRI F	73	112.26	1.54

Table 6: reduced Chi-squared values

This is further quantified in the reduced Chi-squared values in Table 6 which are all larger than 1 (but less than 2) only after we remove those results which have zeta scores greater than 3. The numbers of observations that are removed are generally quite small (around 10-12 or approx. 10% of the data sets), this is entirely consistent with the small number of outliers that are apparent when the individual sample results are assessed. Nonetheless, this does provide direct evidence that there remains some excess scatter in the results above what would be expected given the laboratory quoted errors. Our results suggest that the variability in the tree-ring results is a function of number of rings, with evidence from the reduced Chi-squared statistics, that single tree ring results show reduced variability compared to the tree ring blocks (a reduction of the order of 20%). With regard to the IntCal programme of work, our results include many more laboratories than would necessarily contribute to the master calibration data sets. They do however suggest that the 1 standard deviation envelopes for the curves are too narrow, in each of the time periods we have studied.

3.4 Laboratory benefits of a well characterised reference value

While inter-comparisons are only snap shots in time, one significant benefit from a well designed study using appropriate materials (available in sufficient quantities) is to allow individual laboratories in the future to use well characterised materials as routine reference materials or secondary standards. The FIRI I pine sample is one such reference material which was used in the SUERC laboratory from 2003 till 2009. This was a large wood sample that was power planned to produce wood shavings, and cellulose produced. The cellulose samples were combusted using quartz tubes and two graphite targets were produced from each gas (single combustions). The two targets were then run on random graphite units. In SUERC, the batches of samples are notionally divided into 13 groups of 10 samples, with each group having 3 standards (one oxalic acid II primary standard, one Belfast cellulose secondary standard and either a barley mash or a background standard) and 7 Unknowns. Once the data has been reduced, the average and standard deviation are calculated for the Belfast cellulose, the standard deviation on these values are used to determine the minimum error reported for each batch (Dunbar et al 2016). Table 7 shows the summary of results for 7 years this system was operated as well as the FIRI I intercomparison result. The table shows the within laboratory variability in the sample (where more than 1000 measurements of the sample were made)

Year	2003	2004	2005	2006	2007	2008	2009	FIRI I
n	34	65	159	233	166	285	178	94
Mean (^{14}C BP)	4540	4491	4493	4496	4498	4489	4488	4495
Standard deviation	57	38	45	30	32	30	33	208

Table 7: SUERC summary results for FIRI I

4. Conclusions

The series of wood samples used in the inter-comparisons span a broad range of ages, and include some samples that have been pre-treated before distribution. We have designed the inter-comparisons to include linked samples over time, as well as duplicates both of wood and holo-cellulose. The results provide evidence of the total variability from the potential sources- variation in the samples themselves and differences between laboratories. Such investigations inform on the robustness and repeatability of complex measurements. Importantly, with dendro-dated samples, it is also possible to inform the variability needed when statistically modelling the global calibration curves.

Acknowledgements

The series of wood inter-comparisons would not have been possible without the gracious provision of the samples by many colleagues nor the funding provided by a number of sources including UK research councils (EPSRC and NERC), EU (FP4), NATO, Historic England and Historic Environment Scotland. Finally, a huge thanks to the participating laboratories that have contributed several thousand ^{14}C age measurements over the years.

References

Aitchison T C, Scott E M, Harkness D D, Baxter M S, Cook G T (1990) Report on stage 3 of the International Collaborative Program. Radiocarbon, 32(3) 271-278

Analytical Methods Committee, AMCTB No. 74 (2016) z-Scores and other scores in chemical proficiency testing—their meanings, and some common misconceptions. Anal. Methods, 2016, 8, 5553

Cook G T, Harkness D D, Miller B F, Scott E M, Baxter M S, Aitchison T C (1990) International collaborative study: structuring and sample preparation. Radiocarbon, 32(3), 267-270

Dunbar, E., Cook, G.T., Naysmith, P., Tripney, B.G., Xu, S. (2016) AMS 14C Dating at the Scottish Universities Environmental Research Centre (SUERC) Radiocarbon Laboratory. Radiocarbon 58(1): 9-23.

Harkness, D.D., Cook, G.T., Miller, B.F., Scott, E.M. and Baxter, M.S. (1989). Design and preparation of samples for the international collaborative study. Radiocarbon 31(3): 407-413.

Long A, Kalin R M, 1990. A suggested quality assurance protocol for radiocarbon dating laboratories. Radiocarbon 32(3), 329-334.

Reimer P J, Bard E, Bayliss A, Beck JW, Blackwell P G, Bronk Ramsey C, Buck C E, Cheng H, Edwards R L, Friedrich M, Grootes P M, Guilderson T P, Haflidason H, Hajdas I, Hatté C, Heaton T J, Hoffmann D L, Hogg A G, Hughen K A, Kaiser F K, Kromer B, Manning S W, Niu N, Reimer R W, Richards D A, Scott E M, Southon J R, Staff R A, Turney C S M, van der Plicht J (2013). IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years cal BP. Radiocarbon 55(4), 1869-1887.

Scott, E.M., Aitchison, T.C., Harkness, D.D., Baxter, M.S., Cook, G.T. (1989) An interim progress report on stages 1 and 2 of the international collaborative programs. Radiocarbon, 31, 414-421.

Scott E M, Aitchison T C, Harkness D D, Cook G T, Baxter M S, 1990. An overview of all three stages of the international radiocarbon intercomparison. *Radiocarbon* 32(3), 309-319.

Scott, E.M., Harkness, D.D., Cook, G.T., Aitchison, T.C., Baxter, M.S. (1991) Future quality assurance in 14C dating, Quaternary Proceedings, 1, 1-4.

Scott E M, Harkness D D, Miller B F, Cook G T, Baxter M S, 1992. Announcement of a further international intercomparison exercise. Radiocarbon 34(3), 528-532.

Scott, E M (ed), (2003). The third international radiocarbon inter-comparison (TIRI) and the fourth international radiocarbon inter-comparison (FIRI) 1990-2002: results, analyses, and conclusions, *Radiocarbon*, 45, 135-408.

Scott E M, Cook G T, Naysmith P (2010) The Fifth International Radiocarbon Intercomparison (VIRI): An Assessment of Laboratory Performance in Stage 3. Radiocarbon 52 (3), 859-965

Scott E M, Naysmith P Cook G T (2017) Should Archaeologists Care about 14C Intercomparisons? A Summary Report on SIRI. Radiocarbon 59 (1), pp. 1589-1596

Scott, E.M., Naysmith, P. and Cook, G.T. (2018). Why do we need ^{14}C inter-comparisons? The Glasgow ^{14}C inter-comparison series, a reflection over 30 years. Quaternary Geochronology 43, PP. 72-82.

Stuiver, M., Reimer, P.J., and Braziunas, T.F., (1998), Radiocarbon 40:1127-1151

Thompson M, Ellison S R, Wood R. (2006) The international harmonized protocol for the proficiency testing of analytical chemistry laboratories. Pure Appl. Chem., 78(1), 145–196.