Radiocarbon, Vol 62, Nr 4, 2020, p 873–882 DOI:10.1017/RDC.2020.49

FINDINGS FROM AN IN-DEPTH ANNUAL TREE-RING RADIOCARBON INTERCOMPARISON

L Wacker[1]* • E M Scott[2] • A Bayliss[3] • D Brown[4] • E Bard[5] • S Bollhalder[1] • M Friedrich[6] • M Capano[5] • A Cherkinsky[7] • D Chivall[8] • B J Culleton[9] • M W Dee[10] • R Friedrich[11] • G W L Hodgins[12] • A Hogg[13] • D J Kennett[14] • T D J Knowles[15] • M Kuitems[10] • T E Lange[12] • F Miyake[16] • M-J Nadeau[17] • T Nakamura[16] • J P Naysmith[18] • J Olsen[19] • T Omori[20] • F Petchey[13,21] • B Philippsen[19] • C Bronk Ramsey[8] • G V Ravi Prasad[7] • M Seiler[17] • J Southon[22] • R Staff[18] • T Tuna[6]

[1]Ion Beam Physics, Department of Physics, ETH Zürich, HPK H29, Otto-Stern-Weg 5, 8093 Zürich, Switzerland
[2]School of Mathematics and Statistics, University of Glasgow, Glasgow G12 8QS, UK
[3]Historic England, Cannon Bridge House, 25 Dowgate Hill, London, EC4R 2YA, UK
[4]School of Natural and Built Environment, Queen's University Belfast, Belfast, UK
[5]CEREGE, Aix Marseille University, CNRS, IRD, INRAE, Collège de France, Technopole de l'Arbois BP 80, 13545 Aix en Provence Cedex 4, France
[6]University of Hohenheim, Hohenheim Gardens (772), D-70599 Stuttgart, Germany
[7]Center for Applied Isotope Studies, University of Georgia, 120 Riverbend Road, Athens, GA 30602, USA
[8]Oxford Radiocarbon Accelerator Unit (ORAU), Dyson Perrins Building, Oxford OX1 3QY, UK
[9]Institutes of Energy and the Environment, Pennsylvania State University, University Park, PA 16802, USA
[10]Centre for Isotope Research, University of Groningen, Groningen 9747 AG, The Netherlands
[11]Curt-Engelhorn-Center Archaeometry, C4-8, Mannheim, Germany
[12]University of Arizona AMS Laboratory, Department of Physics, University of Arizona, Tucson, AZ, USA 85721
[13]Radiocarbon Dating Laboratory, University of Waikato, Hamilton, New Zealand
[14]Department of Anthropology, University of California, Santa Barbara, CA 93106, USA
[15]School of Chemistry and Bristol Radiocarbon AMS (BRAMS) Facility, University of Bristol, 43 Woodland Road, Bristol BS8 1UU, UK
[16]Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, Japan
[17]National Laboratory for Age Determination, NTNU University Museum, 7491 Trondheim, Norway
[18]Scottish Universities Environmental Research Centre (SUERC), East Kilbride, South Lanarkshire, Scotland, UK
[19]Aarhus AMS Centre (AARAMS), Department of Physics and Astronomy, Ny Munkegade 120, Aarhus University, DK-8000 Aarhus C, Denmark
[20]Laboratory of Radiocarbon Dating, The University Museum, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo, 113-0033, Japan
[21]ARC Centre of Excellence for Australian Biodiversity and Heritage, College of Arts, Society and Education, James Cook University, Cairns, QLD, Australia
[22]Earth System Science Dept, University of California, Irvine CA 92697-3100, USA

ABSTRACT. The radiocarbon (^{14}C) calibration curve so far contains annually resolved data only for a short period of time. With accelerator mass spectrometry (AMS) matching the precision of decay counting, it is now possible to efficiently produce large datasets of annual resolution for calibration purposes using small amounts of wood. The radiocarbon intercomparison on single-year tree-ring samples presented here is the first to investigate specifically possible offsets between AMS laboratories at high precision. The results show that AMS laboratories are capable of measuring samples of Holocene age with an accuracy and precision that is comparable or even goes beyond what is possible with decay counting, even though they require a thousand times less wood. It also shows that not all AMS laboratories always produce results that are consistent with their stated uncertainties. The long-term benefits of studies of this kind are more accurate radiocarbon measurements with, in the future, better quantified uncertainties.

KEYWORDS: annual tree rings, calibration, high precision, IntCal, radiocarbon intercomparison.

*Corresponding author. Email: wacker@phys.ethz.ch.

MOTIVATION

The tree-ring section of the radiocarbon (^{14}C) calibration curve is overwhelmingly based on measurements made by decay counting in the 1980s and 1990s (all but 56 of the 4314 measurements on tree rings included in IntCal13 were made in this way). At that time, ^{14}C measurements performed by accelerator mass spectrometry (AMS) were not considered of comparable precision to results produced by decay counting (Schmidt et al. 1987; Fifield 2000), in which precision below 2‰ on modern samples was (and still is) possible (Stuiver 1978; Stuiver 1982; Stuiver and Becker 1986; Hogg et al. 2011).

Today, most ^{14}C measurements are made by AMS, largely because much less sample material is required. At least several grams of wood were required for a high-precision measurement with decay counting, but only a few milligrams are required for AMS. While in principle modern AMS systems are capable of measuring at precisions of less than 2‰ on modern samples, it is still an open question if this precision and accuracy can be reached on a routine basis when much smaller samples are processed.

This single tree-ring intercomparison originated as a small experiment to systematically test the reproducibility of AMS measurements on wood samples following discussions at the IntCal meeting in Belfast in 2016. While eight laboratories initially planned to join the exercise, 16 laboratories ultimately took part.

The goal of the exercise was to test the precision and accuracy of the participating laboratories by analyzing and comparing a large number of replicated samples. Three sets of 21 consecutive single tree-ring samples from different time intervals, when no strong changes in ^{14}C concentrations are expected, were used. The high replication should also allow for a meaningful statistical treatment of the results in order to reliably detect any laboratory biases.

The objective of this study was not planned to replace the successful series of international ^{14}C intercomparison exercises (Rozanski et al. 1992; Scott 2003), in which all ^{14}C laboratories are invited to assess their performance on individual samples and a consensus value is derived enabling the samples to subsequently be used as reference materials. A recent study has re-evaluated all AMS measurements on dendrochronologically dated wood samples (which were typically block samples of 20–40 rings) within these previous intercomparisons (Scott et al. 2019). The precision (range of median errors) identified in that study was between 24 to 60 yr for individual measurements on samples with ages between 5000 BP and modern, which is nearly double that of measurements on wood of equivalent age included in IntCal13 (typically between 15 and 30 yr).

Here, our goal was to perform an intercomparison in which high-precision measurements were requested for the specific purpose of testing laboratory performance as part of constructing an accurate, new, annually resolved calibration curve.

SAMPLES AND ANALYSIS

Tree-Ring Samples

Three trees from different time periods within the Holocene were selected for this intercomparison. As the aim of the exercise was to test the laboratories systematically for their ability to produce accurate, high-precision ^{14}C measurements, 21 continuous single tree-ring samples were prepared for each time period. Each 21-yr block was deliberately

selected from a part of the calibration curve where atmospheric ^{14}C varied little (i.e. from a plateau), as this provides a more powerful test of accuracy. The sample size of the raw wood supplied was deliberately limited to typically 30–50 mg. Samples contained earlywood and latewood. Dendrochronological information is given in the supplementary material.

All laboratories obtained 21 contiguous tree-ring samples for the years AD 1730 –1750 (Series H, 220–200 cal BP (AD 1950); Table 2), and AD 280–300 (Series A, 1670–1650 cal BP). For the years 5701–5681 BC (Series R, 7650–7630 cal BP) there was only sufficient material for 9 laboratories, owing to the fact that the exercise was originally planned for 8 laboratories. All laboratories were also supplied with two wood samples containing no detectable ^{14}C that could be employed as a processing blank (brown coal from Reichwalde, Germany and an Oxygen Isotope Stage 5 kauri from New Zealand). While all laboratories processed wood blanks, three laboratories did not use either the supplied lignite or Stage 5 kauri.

Sample Preparation

All laboratories applied a pretreatment with acid and base followed by bleaching. About half of the laboratories started with an overnight base soak (Němec et al. 2010) and one laboratory applied a solvent extraction by Soxhlet (Hoper et al. 1998) as a first step. One third of the laboratories applied an additional α-cellulose step after bleaching (Hoper et al. 1998).

Oxalic Acid 2 (OX2, NIST SRM 4990C) was the most common material used for standard normalization. One laboratory used only Oxalic Acid 1 (OX1, NIST SRM 4990B), while some laboratories used OX1 in addition to OX2.

Radiocarbon Analysis

Instrumentation and Data Reduction

Eight laboratories used a sub-1MV system (5 Micadas 0.2 MV, 3 NEC 0.5 MV) employing the 1+-charge state, while 2 laboratories used a 1 MV (HVEE, 2+-charge state) and 5 used a larger accelerator (3 MV or 5 MV from NEC or HVEE, ≥3+-charge state).

All laboratories using an ETH/Ionplus Micadas system, used the BATS program (Wacker et al. 2010) for data evaluation. The NEC ABC-software was also used frequently, sometimes in combination with the Data Fudger from the Lawrence Livermore AMS laboratory. About 1/3 of all laboratories use their own tools for data reduction (Donahue et al. 1990; Burr et al. 2007), either through coding their own programs or using Excel spreadsheets.

Replication of Data

Typically, each laboratory undertook one measurement per sample. Only one laboratory (Lab-2) systematically undertook duplicate measurements on each annual ring from the outset. A few laboratories repeated either specific series or some samples to test their internal variability. All repeats were performed on the same cellulose extraction with the exception of 4 rings of series R and A for Lab-2, for which the cellulose extraction was performed twice. Lab-9 always combined two consecutive tree rings so that enough material was available to survive the more rigorous α-cellulose treatment.

One laboratory (Lab-6) decided to retract their data for series H due to possible detector instabilities that became obvious after initial submission. Laboratory 14 submitted two independent measurement series of which only the second has been evaluated and is reported here (on the suggestion of the laboratory). Because of memory effects from a dirty ion source, the blank was 4-times higher than expected in a first run and consequently a second run was performed with a clean source. Interestingly, the dirty ion source resulted in ^{14}C ages that are consistently 20 yr older.

Ultimately, more than 950 measurements were performed in total (H-series: 396, A-series: 353, R-series 203). Each ring was measured between 10 (R-series) and 20 times (H-series).

Blank Correction

Ten laboratories reported on blanks used for their data reduction. Most laboratories (8) relied only on full processing blanks, while one laboratory used both combustion and processing blanks, and one only used combustion blanks. The average blank used for the correction was $F^{14}C = 0.0019$ (scatter: 0.0005). No consistent information was obtained about the estimated reproducibility of the blanks by individual laboratories but would be advantageous when sample ages are closer to the background (>10,000 BP).

Uncertainties by Laboratory

The laboratories followed very different approaches for estimating uncertainties. While many laboratories calculate their uncertainties on the larger of the counting statistic uncertainty or the internal variability of a sample measurement (variability of submeasurements of a single sample over time), other laboratories base their calculations on counting statistics where they add, in quadrature, an additional uncertainty (of typically 1–1.5‰), before adding an uncertainty for standard normalization and blank correction. While the first approach results in a relatively large variation in uncertainty, even within the same set of measurements (see also min and max uncertainties in Table 1), the latter provides less-variable uncertainties.

Statistical Analysis

In a first step, the error weighted mean for replicate measurements by an individual laboratory on each annual ring was calculated, assuming a systematic uncertainty of 1‰ from sample preparation (Sookdeo et al. 2019 in this issue) that is already included in the uncertainties given by the laboratories. The arithmetic mean ($\overline{x}_{ring}$) and the median for each ring was determined, including the results of all 16 laboratories (AL). The mean is considered to be the consensus value for the individual tree-ring samples with the error estimated from the variance within the set of results. Thus, this error does not take account of the individual errors quoted by the laboratories. Statistical tests were typically performed on single measurements from specified laboratories against the consensus values (AL).

The distribution of a selection of laboratories (SL), which showed both (a) average uncertainties given by the laboratory of less than 20 yr on series H and A and (b) least variable offsets from the mean (Lab-1, 2, 5, 7, 8, 9, 14, 15–see also Discussion), was additionally compared to the distribution of AL. The offset (off_{lab}) is the mean deviation of all n measurements of a laboratory (x_{ring_i}) to the mean ($\overline{x}_{ring}$):

Table 1 The arithmetic means with its uncertainty and the median for all tree rings of the three series is given.

H-series				A-series				R-series			
Ring #	Age (AD)	^{14}C age (BP)	Median ^{14}C age	Ring #	Age (AD)	^{14}C age (BP)	Median ^{14}C age	Ring #	Age (BC)	^{14}C age (BP)	Median ^{14}C age
1	1730	153.2 ± 8.8	151.4	1	280	1728.1 ± 5.5	1718.0	1	5701	6814.9 ± 15.5	6805.7
2	1731	159.7 ± 5.0	161.0	2	281	1726.7 ± 4.4	1732.7	2	5700	6791.7 ± 9.0	6783.3
3	1732	155.7 ± 4.9	155.9	3	282	1728.8 ± 6.0	1734.8	3	5699	6820.5 ± 7.3	6813.4
4	1733	166.5 ± 7.0	164.5	4	283	1732.5 ± 6.0	1735.4	4	5698	6821.8 ± 11.3	6814.7
5	1734	156.7 ± 4.1	161.2	5	284	1738.9 ± 4.2	1741.7	5	5697	6843.0 ± 18.8	6839.6
6	1735	161.0 ± 5.2	156.5	6	285	1745.4 ± 7.3	1742.5	6	5696	6813.5 ± 7.7	6813.0
7	1736	156.0 ± 6.3	159.2	7	286	1748.4 ± 7.6	1756.2	7	5695	6818.3 ± 9.2	6805.8
8	1737	159.2 ± 5.2	159.5	8	287	1749.1 ± 3.8	1746.3	8	5694	6808.6 ± 5.4	6807.4
9	1738	161.6 ± 5.6	164.7	9	288	1746.1 ± 4.2	1745.4	9	5693	6837.0 ± 12.8	6836.2
10	1739	171.6 ± 4.4	169.0	10	289	1750.4 ± 7.1	1739.6	10	5692	6833.4 ± 6.9	6834.1
11	1740	174.5 ± 7.4	170.1	11	290	1744.1 ± 8.0	1741.7	11	5691	6833.2 ± 7.7	6836.7
12	1741	182.5 ± 7.9	174.2	12	291	1739.3 ± 7.5	1739.9	12	5690	6847.7 ± 13.8	6842.9
13	1742	187.7 ± 4.5	189.1	13	292	1752.5 ± 7.6	1752.6	13	5689	6853.9 ± 15.3	6854.7
14	1743	184.8 ± 4.3	188.9	14	293	1755.1 ± 5.4	1761.5	14	5688	6847.2 ± 9.1	6838.1
15	1744	174.3 ± 6.4	177.0	15	294	1756.5 ± 4.7	1760.1	15	5687	6847.1 ± 13.8	6859.4
16	1745	189.8 ± 8.3	191.4	16	295	1758.4 ± 5.9	1760.4	16	5686	6823.1 ± 11.1	6830.6
17	1746	180.7 ± 3.9	178.4	17	296	1760.1 ± 6.2	1759.7	17	5685	6822.4 ± 11.6	6828.4
18	1747	188.6 ± 5.7	189.3	18	297	1752.4 ± 5.6	1755.4	18	5684	6822.6 ± 6.4	6817.9
19	1748	189.6 ± 6.3	190.2	19	298	1753.0 ± 5.1	1755.4	19	5683	6825.2 ± 12.8	6838.0
20	1749	195.8 ± 5.3	195.0	20	299	1754.8 ± 6.8	1749.2	20	5682	6845.5 ± 6.2	6851.9
21	1750	184.0 ± 6.5	181.2	21	300	1752.7 ± 6.9	1751.2	21	5681	6830.2 ± 12.2	6845.8

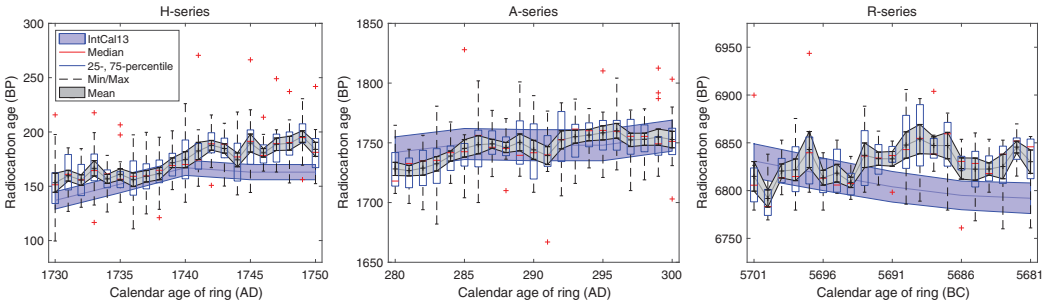

Figure 1 Box plot of the measurements of all laboratories for the individual series and the arithmetic means (black, gray) compared with IntCal13. The red data points are outside 1.5-times the interquartile range (indicated by min/max). (Please see electronic version for color figures.)

$$off_{lab} = \sum_{ring,i}^{n} \frac{(\overline{x}_{ring} - x_{ring_i})}{n}$$

The mean uncertainties produced by the laboratories were independently estimated for each series by comparing the measurements of a laboratory with the consensus values of all laboratories after subtraction of the laboratory offset:

$$\sigma_{lab_{series}} = \sqrt{\sum_{ring,i}^{n} \frac{(\overline{x}_{ring} - (x_{ring_i} - off_{lab}))^2}{n}}$$

The errors given by a laboratory (unc_{ring_i}) are finally compared with the observed deviations of the ring arithmetic mean value adjusted for the laboratory offset:

$$\chi^2_{red_{lab_{series}}} = \frac{\sum_{ring,i}^{n} \frac{(\overline{x}_{ring} - (x_{ring_i} - off_{lab}))^2}{unc_{ring_i}^2}}{n}$$

RESULTS AND DISCUSSION

Sample Mean and Comparison with IntCal13

The median and the arithmetic mean agree well with each other (Table 1), indicating a symmetrical distribution of data from all laboratories.

The mean, as well as the median, of all individual tree-ring measurements was calculated and compared with IntCal13, as shown with box plots in Figure 1. Series H is significantly older than the IntCal13 data. The offset is on average 16 ± 4 ^{14}C yr (in the following simply given as years), although this offset varies over the 21 yr supplied. While the mean offset of series A is –2 ± 6 yr, in good agreement with IntCal13, the offset of series R is 23 ± 8 yr, which is again statistically significant. We should note however, that IntCal13 is based on just a few decadal measurements in these time periods and consequently, offsets to IntCal13 cannot be determined as precisely. It may show that the interpolation between those few measured

samples calculated for IntCal13 (IntCal13 provides data points for every 5th year) does not reflect the actual pattern of atmospheric ^{14}C closely.

Laboratory Offsets

For each series, the mean laboratory offset relative to the overall mean and median is given in Table 2. As expected from the good agreement of the median and the mean, the offsets relative to the mean and median are all very similar. As a rule of thumb, deviations of less than 10 yr are acceptable within 2 σ, due to counting statistics and systematic effects of standard normalization. While about half of the laboratories seem to be consistently within ±10 yr (Labs-1, 2, 5, 7, 8, 9, 14, and 15), three laboratories have offsets of 10–20 yr (Labs-11, 13, 16), while other laboratories show larger and at the same time more variable offsets of up to 20–30 yr (Labs-3, 4, 6, 10, and 12). Over all, the measured offsets seem to be comparable to those observed on high-precision measurements (IntCal98) in the 1990s by decay counting (Stuiver et al. 1998), in which laboratory offsets, primarily on measurements on decadal samples, were typically between 10 and 20 yr. However, today many more laboratories are able to produce data comparable to what was available in the 1990s. If only the 8 laboratories with the most consistent datasets (SL) are considered, a better constrained consensus value can be calculated.

Laboratory offsets seem to be distributed symmetrically, with large deviations to both younger and older ages. Laboratories with large offsets seem to be either consistently younger or consistently older for all series, pointing towards intrinsic systematic causes for the offsets. This observation needs to be investigated further by the laboratories concerned, as no obvious correlation is observed between these offsets and the information supplied about pretreatment protocols, blank subtraction, and standard normalization.

Estimation of Uncertainties

In order to better understand the accuracy of single measurements, their relative deviations from their corresponding mean are plotted as a histogram in Figure 2. The distribution is Gaussian-like in shape, and has a spread of about 21 yr for series H and A. This is clearly larger than the median of 17.5 yr for the quoted uncertainties. A reduced chi-square-test, comparing the deviations of the individual laboratories to the mean with their quoted uncertainties, shows that there is evidence of over-dispersion in comparison to what would be expected given the quoted uncertainties ($\chi^2_{red} \geq 1.6$). This indicates that some laboratories are underestimating their uncertainties.

In contrast, the selection of laboratories (SL) with the smallest and least variable offsets show significantly narrower distributions (σ_{SL}) for all series. The quoted uncertainties provided by these laboratories are also consistent with the observed deviations (although we are still comparing with the consensus value calculated from all of the data). This is indicated in Figure 2, with χ^2_{red} values for these laboratories of ~1.0 for all series.

The distributions of the SL show that AMS measurements from different laboratories can reproduce well within 2‰ (16 yr) for modern samples. The distribution is comparable to the one obtained for single-year data produced by a single laboratory at the University of Washington (QL) in the 1980s with high-precision gas proportional counting measurements (σ_{QL}=14.4) (Stuiver et al. 1998). However, here the data stem from 8 different laboratories, which is a significant advancement. The SL also get close to reproducing within 2‰

Table 2 The performance of the laboratories relative to the mean and the median is given. Offsets and uncertainties are given in ^{14}C yr (BP). The independently estimated uncertainty is calculated from the offset corrected deviations from the mean respectively the median. The χ^2_{red} is calculated from the estimated uncertainties relative to the quoted uncertainties.

H-Series	Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7	Lab 8	Lab 9	Lab 10	Lab 11	Lab 12	Lab 13	Lab 14	Lab 15	Lab 16
Dendro Record	HFD-C01	HFD-C02	HFD-C03	HFD-C06	HFD-C07	HFD-C08	HFD-C09	HFD-C10	GBD-A101	BRT-A17	GBD-A102	GBD-A103	GBD-A105	GBD-A106	GBD-A107	GBD-A108
offset (mean)	3.8	7.6	-8.8	18.2	-0.1		-0.3	2.1	-0.1	-25.2	-14.0	0.4	13.7	-1.8	8.2	-14.3
offset (median)	4.0	6.8	-9.2	20.3	0.5		0.5	1.4	0.0	-26.7	-13.4	0.3	14.7	-1.3	8.2	-14.5
estimated unc.	20.4	20.5	24.3	25.7	10.9		10.4	15.9	15.1	16.2	22.5	22.9	22.6	12.2	17.6	17.1
given mean unc.	15.9	18.2	18.4	25.5	16.0		15.2	20.2	14.6	16.9	20.7	15.1	29.4	14.0	11.3	21.1
min given unc.	15.3	18.0	15.0	21.0	15.4		15.0	20.0	13.0	15.0	19.0	13.5	27.6	13.9	11.0	21.0
max given unc.	16.9	20.0	21.0	29.0	17.0		20.0	22.0	15.0	20.0	26.0	17.4	33.3	14.0	15.0	22.0
χ^2_{red} (mean)	1.7	1.4	1.8	1.4	0.4		0.5	0.6	0.9	3.4	1.6	2.1	1.1	0.7	2.8	1.1
χ^2_{red} (median)	1.6	1.4	1.8	1.6	0.5		0.5	0.6	0.9	3.4	1.6	2.2	1.1	0.8	2.7	1.2
# of measurements	35	42	62	29	21		24	21	10	21	21	21	20	19	22	28
A-Series	Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7	Lab 8	Lab 9	Lab 10	Lab 11	Lab 12		Lab 14	Lab 15	Lab 16
offset (mean)	-3.0	3.0	3.2	24.2	11.0	-30.5	6.1	7.8	-6.3	-14.2	-10.8	19.7		-2.7	-0.2	-5.0
offset (median)	-3.0	3.7	2.8	26.1	11.4	-32.8	6.3	7.9	-8.3	-14.3	-10.0	22.8		-3.4	-1.2	-4.4
estimated unc.	11.6	19.4	20.4	30.7	11.8	14.1	12.4	18.6	12.2	15.0	24.7	17.1		10.2	14.5	13.8
given mean unc.	13.2	18.5	19.2	27.4	15.9	15.8	15.9	23.2	14.7	20.0	26.1	17.2		14.7	13.3	20.4
min given unc.	13.1	18.0	15.0	23.5	15.2	13.0	15.0	21.0	14.0	20.0	24.0	12.4		14.5	11.0	20.0
max given unc.	13.5	19.0	21.0	31.0	16.4	21.0	20.0	37.0	17.0	20.0	32.0	23.8		14.9	16.0	28.0
χ^2_{red} (mean)	0.8	1.1	1.0	2.0	1.0	4.7	0.7	0.8	0.8	1.0	1.0	1.8		0.5	1.2	0.5
χ^2_{red} (median)	0.9	1.2	1.1	2.1	1.0	5.0	0.8	0.8	0.7	1.1	0.9	1.9		0.6	1.6	0.6
# of measurements	21	46	42	28	21	21	23	21	10	21	21	15		19	21	23
R-Series	Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7	Lab 8	Lab 9							
offset (mean)	-1.7	2.2	28.0	-1.4	17.2	-24.7	-0.5	2.0	-4.4							
offset (median)	-2.0	2.1	28.4	-1.0	18.2	-28.7	-2.5	2.5	-6.7							
estimated unc.	19.1	22.6	20.2	46.0	19.1	22.6	21.3	23.8	18.8							
given mean unc.	16.4	23.0	21.0	38.9	24.4	21.7	19.5	28.0	15.4							
min given unc.	16.2	22.0	21.0	33.0	23.6	17.0	15.0	20.0	15.0							
max given unc.	16.9	24.0	21.0	45.0	25.5	36.0	20.0	36.0	16.0							
χ^2_{red} (mean)	1.3	0.9	2.7	1.3	1.1	2.4	1.2	0.8	1.4							
χ^2_{red} (median)	1.1	0.9	2.9	1.5	1.0	2.9	1.3	0.7	1.2							
# of measurements	24	46	18	21	21	21	21	21	10							

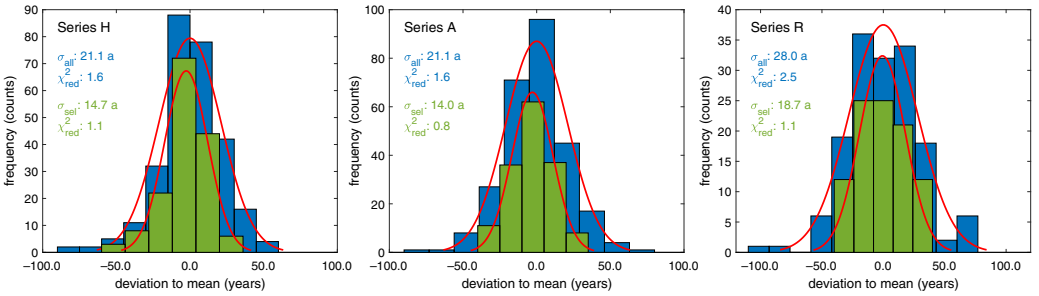

Figure 2 Distributions of deviations of all measurements of all laboratories (blue) and a selection of laboratories (green) around the means of the individual series. Positive deviations correspond to older ages. (Please see electronic version for color figures.)

(see Figure 2, R series) for samples over 7000 yr old. Finally, it is also worth noting that three laboratories (Lab-5, Lab-7, and Lab-8) seem to be over-estimating their uncertainties, Lab-8 for all three series, and Lab-5 and Lab-7 for the youngest samples (H- and A-series), although not for the older (R-series). This seems to be in contrast to the majority of the other laboratories that tend to underestimate the uncertainties of the younger samples.

CONCLUSIONS

The ^{14}C intercomparison on single-year tree-ring samples reported here is the first to specifically investigate possible offsets between AMS laboratories at high precision. As the study is based on a large number of measurements, it also allows us to estimate the accuracy of measurements independent from the errors quoted by the laboratories. Consequently, this type of intercomparison exercise will be an important instrument for future progress in creating new, more structured, more precise, and more accurate ^{14}C calibration curves.

The results show that AMS laboratories today are capable of measuring samples of Holocene age with an accuracy that is comparable or even goes beyond what is possible with decay counting (even though they require a thousand times less wood). But it also shows that not all AMS laboratories always produce results that are consistent with their stated uncertainties. Consequently, this exercise has proven to be a valuable tool in identifying inter-laboratory offsets. The long-term benefits of studies of this kind are more accurate ^{14}C measurements with, in the future, better quantified uncertainties.

ACKNOWLEDGMENTS

M. Kuitems and M. W. Dee were able to participate in this intercomparison thanks to the support of a European Research Council grant (ECHOES, 714679). AixMICADAS and its operation at CEREGE are funded by the Collège de France, the EQUIPEX ASTER-CEREGE and the ANR project CARBOTRYDH (PI E.B.).

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/RDC.2020.49

REFERENCES

Burr GS, Donahue DJ, Tang Y, Beck JW, McHargue L, Biddulph D, Cruz R, Jull AJT. 2007. Error analysis at the NSF-Arizona AMS facility. Nuclear Instruments & Methods in Physics Research B 259(1):149–153.

Donahue DJ, Linick TW, Jull AJT. 1990. Isotope-ratio and background corrections for accelerator mass-spectrometry radiocarbon measurements. Radiocarbon 32(2):135–142.

Fifield LK. 2000. Advances in accelerator mass spectrometry. Nuclear Instruments & Methods in Physics Research B 172:134–143.

Hogg A, Palmer J, Boswijk G, Turney C. 2011. High-precision radiocarbon measurements of tree-ring dated wood from New Zealand: 195 BC–AD 995. Radiocarbon 53(3):529–542.

Hoper ST, McCormac FG, Hogg AG, Higham TFG, Head MJ. 1998. Evaluation of wood pretreatments on oak and cedar. Radiocarbon 40(1):45–50.

Němec N, Wacker L, Hajdas I, Gäggeler H. 2010. Alternative methods for cellulose preparation for AMS measurement. Radiocarbon 52(3):1358–1370.

Rozanski K, Stichler W, Gonfiantini R, Scott EM, Beukens RP, Kromer B, van der Plicht J. 1992. The IAEA C-14 Inter-comparison exercise 1990. Radiocarbon 34(3):506–519.

Schmidt FH, Balsley DR, Leach DD. 1987. Early expectations of AMS – Greater ages and tiny fractions – One failure – One success. Nuclear Instruments & Methods in Physics Research B 29(1–2):97–99.

Scott EM. 2003. The Fourth International Radiocarbon Intercomparison (FIRI). Radiocarbon 45(2): 135–150.

Scott EM, Cook GT, Naysmith P, Staff RA. 2019. Learning from the wood samples in ICS, TIRI, FIRI, VIRI, and SIRI. Radiocarbon 61(5): 1293–1304.

Sookdeo A, Kromer B, Buentgen U, Friedrich M, Friedrich R, Helle G, Pauly M, Nievergelt D, Reinig F, Treydte K, Synal HA, Wacker L. 2019. Quality Dating: A well-defined protocol implemented at ETH for high-precision ^{14}C dates tested on Late Glacial wood. Radiocarbon 62. This issue. doi: 10.1017/RDC.2019.132.

Stuiver M. 1978. The ultimate precision of ^{14}C dating is determined only by counting statistics. In: Gove HE, editor. Proc. 1st Conf. on Radiocarbon Dating with Accelerators. Rochester (NY): University of Rochester. p. 353.

Stuiver M. 1982. A high-precision calibration of the AD radiocarbon time scale. Radiocarbon 24(1):1–26.

Stuiver M, Becker B. 1986. High-precision decadal calibration of the radiocarbon time scale, AD 1950–2500 BC. Radiocarbon 28(2B): 863–910.

Stuiver M, Reimer PJ, Bard E, Beck JW, Burr GS, Hughen KA, Kromer B, McCormac G, van der Plicht J, Spurk M. 1998. INTCAL98 radiocarbon age calibration, 24,000–0 cal BP. Radiocarbon 40(3):1041–1083.

Wacker L, Christl M, Synal HA. 2010. BATS: A new tool for AMS data reduction. Nuclear Instruments and Methods in Physics Research B 268(7–8):976–979.