Radiocarbon, Vol 57, Nr 3, 2015, p 377–388 DOI: 10.2458/azu_rc.57.18390

HIGH-PRECISION ATMOSPHERIC $^{14}CO_2$ MEASUREMENT AT THE RAFTER RADIOCARBON LABORATORY

Jocelyn C Turnbull[1,2,3] • Albert Zondervan[1] • Johannes Kaiser[1] • Margaret Norris[1] • Jenny Dahl[1] • Troy Baisden[1] • Scott Lehman[2]

ABSTRACT. This article describes a new capability for high-precision ^{14}C measurement of CO_2 from air at the Rafter Radiocarbon Laboratory, GNS Science, New Zealand. We evaluate the short-term within-wheel repeatability and long-term between-wheel repeatability from measurements of multiple aliquots of control materials sourced from whole air. Samples are typically measured to 650,000 ^{14}C counts, providing a nominal accelerator mass spectrometry (AMS) statistical uncertainty of 1.3‰. No additional uncertainty is required to explain the within-wheel variability. An additional uncertainty factor is needed to explain the long-term repeatability spanning multiple measurement wheels, bringing the overall repeatability to 1.8‰, comparable to other laboratories measuring air materials to high precision. This additional uncertainty factor appears to be due to variability in the measured ^{14}C content of OxI primary standard targets, likely from the combustion process. We observe an offset of 1.4‰ in our samples relative to those measured by the University of Colorado INSTAAR, comparable to interlaboratory offsets observed in recent intercomparison exercises.

INTRODUCTION

The radiocarbon content of atmospheric CO_2 ($^{14}CO_2$) has been measured at sites around the world since 1954 (e.g. Rafter and Fergusson 1959). It has long been used to trace the injection of anthropogenically produced ^{14}C from atmospheric nuclear weapons testing into the atmosphere and its subsequent movement throughout the carbon cycle reservoirs. In recent years, $^{14}CO_2$ has become the tracer of choice to detect and quantify the addition of fossil-fuel CO_2 into the atmosphere (e.g. Levin et al. 2003). Atmospheric $^{14}CO_2$ measurements can also be used to investigate ocean carbon exchange processes (e.g. Graven et al. 2012). Measurements of $^{14}CO_2$ are a powerful tool for understanding global, regional, and local carbon cycle processes, yet the signals of interest are often small, so that the quality of the $^{14}CO_2$ measurements is critical. The World Meteorological Organization (WMO) recommends measurement precision of 3‰ or better in $\Delta^{14}C$ for atmospheric $^{14}CO_2$ measurements to be useful and an ultimate goal of 0.5‰, although it is recognized that 0.5‰ is beyond current measurement capabilities (Tans and Zellweger 2014). To date, several laboratories have documented better than 2‰ long-term repeatability on single sample measurements (Graven et al. 2007; Turnbull et al. 2007; Lehman et al. 2013).

This article describes the high-precision atmospheric $^{14}CO_2$ measurement capability at the Rafter Radiocarbon Laboratory at GNS Science, New Zealand. We discuss the two different methods we commonly use for CO_2 collection and subsequent extraction. We describe our recently upgraded graphitization system and detail our measurement protocols for samples requiring high precision. We use replicate measurements of CO_2 from a number of different control materials, all derived from whole air, to examine mean values and offsets between different standardization methods within our laboratory, as well as interlaboratory offsets based on an ongoing intercomparison with the University of Colorado INSTAAR. We then examine short-term within-wheel and long-term repeatability of our measurements, and the sources of uncertainty that contribute to these.

METHODS

CO_2 from air is collected either by *in situ* absorption of CO_2 into sodium hydroxide (NaOH) solution or by collection of whole air into flasks or pressurized cylinders. In addition, control materials, primary standards, and process blanks are routinely analyzed. In this section, we describe the collection methods and protocols for extracting and purifying the CO_2 in the laboratory.

1. Rafter Radiocarbon Laboratory, GNS Science, Lower Hutt, New Zealand.
2. Corresponding author. Email: j.turnbull@gns.cri.nz.
3. University of Colorado at Boulder, Boulder, CO, USA.

Proceedings of the 1st International *Radiocarbon in the Environment* Conference
18–22 August 2014, Queen's University Belfast, Belfast, Northern Ireland, UK
Edited by Evelyn Keaveney and Paula Reimer

Sodium Hydroxide Absorption

In this method, CO_2 is absorbed into a solution of NaOH, typically 0.5 to 1.0 molarity. The sodium hydroxide solution is exposed to air for a set period, then can be stored indefinitely in a closed vessel. This method allows collection of large amounts of CO_2, which is essential for radiometric counting of ^{14}C, and gives the opportunity for multiple replicate measurements of a single sample by accelerator mass spectrometry (AMS). CO_2 is very readily absorbed into the solution, so our methods are designed to ensure that the initial solution is free of CO_2, and that CO_2 contamination does not occur during the laboratory extraction. Static absorption into a bottle or tray results in significant isotopic fractionation of CO_2 relative to the atmosphere, which we account for by measuring the $\delta^{13}C$ of the absorbed CO_2 by isotope ratio mass spectrometry (IRMS) and, in recent samples, also in-line by AMS (Manning et al. 1990; Currie et al. 2009). Although other researchers have shown that systems wherein the air is pumped through the NaOH solution can reduce fractionation and allow more control over the sampling period (e.g. Levin and Kromer 1997; Van der Laan 2010), for consistency we continue to use the static method for our long-term atmospheric records.

CO_2 is evolved in the laboratory by acidifying the NaOH solution with phosphoric acid and cryogenically trapping the evolved CO_2 (Currie et al. 2009). For this study, we used four different authentic samples collected using static NaOH absorption at Baring Head, New Zealand, between 2000 and 2010, as well as a NaOH absorption sample collected in Wellington city for use as a control material (NIWAair). For each sample, a single CO_2 evolution was performed, and the resulting CO_2 (containing 5–50 mg C) was split under equilibrium conditions into multiple aliquots of 0.5–1.0 mg C. Thus, in this study, the samples derived from NaOH collection test the variability of graphitization and AMS measurement, but do not address possible variability related to sample collection.

Whole-Air Flasks and Cylinders

Glass flask samples of 1–4 L of whole air are collected either by grab sampling or 1-hr integrated air averages (e.g. Conway et al. 2011; Turnbull et al. 2012). It is believed that whole air can be stored for some months without significant alteration of the $^{14}CO_2$ content. The CO_2 is then extracted by cryogenic purification, obtaining 0.2–0.8 mg C from each sample. In this study, we use pressurized whole-air cylinders as control samples. Control aliquots are extracted individually, and extraction times are controlled so as to match the size of authentic flask samples. Thus, any variability in the cryogenic extraction procedure is included in the uncertainty analysis for these control materials, as well as contributions from graphitization and AMS measurement.

CO_2 from flasks or control cylinders is obtained by manual cryogenic extraction at Rafter, using methods based on those described by Turnbull et al. (2007). The CO_2 aliquot is stored in a valved bottle for no more than 5 days before graphitization. Occasionally, when longer CO_2 storage is required, the CO_2 aliquots are stored in breakseal tubes prior to graphitization.

Four control cylinders are included in our analysis here. Rafter holds two of these; BHDamb2013 is natural air collected at Baring Head, New Zealand, in 2013, and BHDspike2013 is natural air spiked with a small amount of ^{14}C-free CO_2 to give a lower-than-ambient $\Delta^{14}C$ value. In addition, we performed measurements on CO_2 extraction aliquots from two control cylinders, NWT3 and NWT4, maintained and used at the University of Colorado INSTAAR since 2009 (Lehman et al. 2013). NWT3 is natural air collected at Niwot Ridge, Colorado, in 2009 and NWT4 is natural air collected at the same time and spiked with ^{14}C-free CO_2. CO_2 aliquots for this study were prepared by automated cryogenic extraction at University of Colorado INSTAAR (Turnbull et al. 2010) and shipped to Rafter in breakseal tubes for graphitization and AMS measurement.

Oxalic Acid I Primary Standard

Rafter uses oxalic acid I (OxI) as the primary measurement standard. CO_2 aliquots are prepared in two ways. In the first method, sufficient OxI for 5 CO_2 aliquots is prepared by sealed tube (ST) combustion with CuO and Ag wire at 900°C. The resulting CO_2 is cryogenically purified and split into 5 aliquots under equilibrium conditions. Off-line IRMS $\delta^{13}C$ measurement is performed on 1 of the 5 aliquots as an additional quality control measure. In the second method, an elemental analyzer (EA) is used for combustion, followed by automated cryogenic purification and CO_2 collection (Baisden et al. 2013). In this case, each CO_2 aliquot is derived from an individual combustion (i.e. not split); 1% of the CO_2 gas is used for $\delta^{13}C$ measurement via continuous-flow IRMS, in-line with the EA. Typically 6 but up to 12 CO_2 aliquots are prepared from a single EA run. In either case, OxI CO_2 from at least two different ST parent combustions or from two different EA runs are included in each AMS measurement wheel.

Blank Materials

We use several nominally ^{14}C-free materials to determine the process blank. SyntheticDeadAir2 and DeadAir2013 are cylinders of CO_2-free air spiked with ^{14}C-free CO_2 to ~400 ppm, and are housed at INSTAAR and Rafter, respectively. Aliquots of CO_2 from these cylinders are extracted in the same manner as for the other air cylinders. These represent the overall blank for air materials. A cylinder of ^{14}C-free CO_2 obtained from the Kapuni natural gas field in New Zealand ("Kapuni CO_2") is used to separately diagnose the blank contribution from graphitization and AMS measurement alone.

GRAPHITIZATION

At Rafter, CO_2 is reduced to graphite using the hydrogen method. Our current Rafter Graphitization 20 reactor system (RG20) was installed in 2012. RG20 replaces an older graphitization system that had been in use since the 1980s (Lowe and Judd 1987) and includes a number of changes to improve data quality, capacity, and efficiency. RG20 is semiautomated, allowing real-time monitoring and recording of reaction progress and automated furnace switch off when reaction completion criteria are met. CO_2 transfer on this system is performed manually.

Initially, 1.6 to 2.0 mg of iron catalyst is prepared by reducing Fe_2O_3 (Sigma Aldrich, 99.999% purity) to pure iron (Fe) by reaction with ~1300 mbar of H_2 gas at 400°C. The resultant water is removed by freezing with a thermoelectric cooler at –18°C. The Fe_2O_3 reduction is monitored and recorded using a pressure transducer. Reduction typically takes 45 min but is continued for at least 1 hr or until the reaction is complete. The reactor area is then allowed to return to room temperature and the water is pumped away.

The sample CO_2 aliquot is introduced into the reactor and H_2 gas added to a pressure of ~2.3 times the CO_2 pressure to ensure sufficient H_2 for complete graphitization. The reaction is performed at 550°C, and water is removed by thermoelectric cooler at –18°C. The CO_2 reduction typically takes 2 hr, and heating is continued for at least 2.5 hr or until the reaction is complete, as determined by a stable pressure measured inside the reactor. Pressure yield is also monitored to ensure complete reaction occurred. The resulting graphite is stored in cleaned plastic vials for a maximum of 1 month and then pressed into aluminum targets (cathode holders) just prior to AMS measurement.

AMS MEASUREMENT AND DATA REDUCTION

For the ^{14}C measurement, we use the Extended Compact Accelerator Mass Spectrometry system (XCAMS) installed at Rafter in 2010 (Baisden et al. 2013; Zondervan et al. 2015). Each ^{14}C measurement wheel contains up to 40 graphite targets. For high-precision atmospheric measurements,

the wheel includes two tuning targets (one modern, one Kapuni CO_2 blank), eight OxI primary standards, three to eight control samples derived from whole air, one ^{14}C-free process blank (either SyntheticDeadAir2 or DeadAir2013), and the remainder are authentic samples. The wheel is rotated through all targets multiple times, providing multiple 2-min "exposures" for each target. Typically, 25 or more exposures are needed to acquire ~650,000 ^{14}C counts on each near-modern target. In a few cases, we have acquired 1,200,000 ^{14}C counts per target.

Our in-house CALAMS processing software is used to determine the ^{14}C content of each sample, using the quasi-simultaneous measurements of all three carbon isotopes provided by XCAMS (Zondervan et al. 2015). Any mass-dependent isotopic fractionation during processing and measurement is thus accounted for in the CALAMS data processing. Process blank correction is applied following Donahue et al. (1990) using the process blank target included in the same measurement wheel. We monitor the performance of both blank materials through time, and the correction to our modern materials is negligible. Results are reported as $F^{14}C$ (Reimer et al. 2004) corrected for process blank and normalized to $\delta^{13}C$ of –25‰, and/or as $\Delta^{14}C$ which is derived from $F^{14}C$ and decay corrected to the date of collection (Stuiver and Polach 1977).

RESULTS AND DISCUSSION

We examine mean values and biases between methods, and the contributions to the overall uncertainty, using repeated measurements of air materials and of OxI primary standards (Figure 1). We use all of our high-precision atmospheric measurement wheels in this analysis, compromising 42 individual wheels measured from 2011 to 2014. A total of 265 targets from 9 different air materials and 350 OxI targets are included.

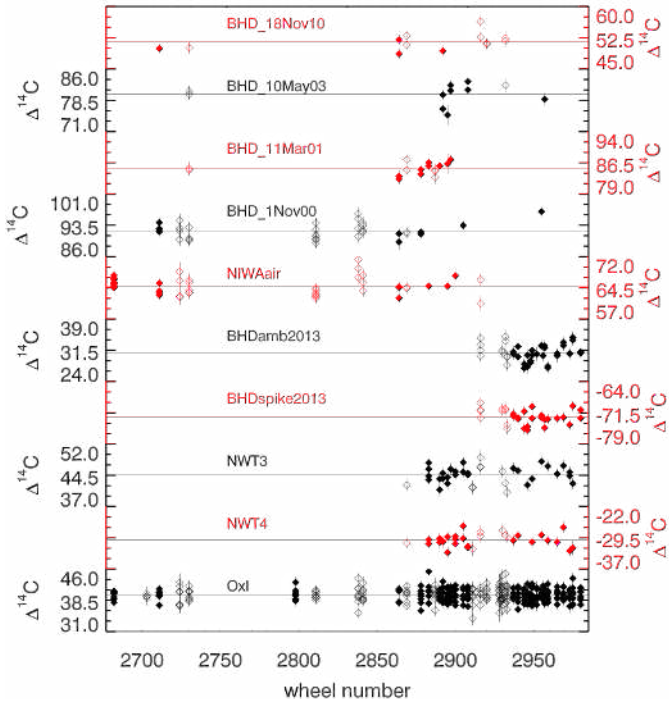

Figure 1 Raw $\Delta^{14}C$ and σ_{AMS} in units of ‰ for all high-precision atmospheric control samples and OxI primary standard measured at Rafter since 2011. Long-term mean $\Delta^{14}C$ values are indicated by lines. For OxI primary standard, the mean value within each wheel is necessarily the same as the long-term mean. Open and closed symbols indicate that the OxI primary standards used in that wheel were prepared by sealed-tube (ST) or elemental analyzer (EA) combustion, respectively. The ST offset described in the text has not been applied in this figure.

Long-Term Mean $\Delta^{14}C$ and Method Biases

Observed $\Delta^{14}C$ values and their raw AMS statistical uncertainties are shown in Figure 1, and mean

values for each material are listed in Table 1. Mean $\Delta^{14}C$ is on average 1.5‰ higher for wheels where the OxI primary standards were prepared by ST than for EA wheels (Figure 1, Table 1). The standard error we calculate in this offset (0.2‰) assumes that the offset is consistent across all materials and all wheels, so if there is variability through time or by material, the error in the offset is underestimated. We also observe more scatter in the results for ST wheels, as evidenced by larger standard deviations for ST wheels in Table 1. Some wheels prepared at Rafter for routine (not high precision) AMS measurement have included a mixture of OxI targets prepared by EA and ST. For that data set, when we treat the EA OxI targets as primary standards, the ST OxI is higher by about 1.5‰, but varying through time (data not shown). NWT3 and NWT4 are measured at both Rafter and INSTAAR and we observe that our EA mean $\Delta^{14}C$ is closer to the INSTAAR mean $\Delta^{14}C$ for these materials (Table 1). Although we cannot establish which combustion method is "correct" in an absolute sense, the EA method appears to be more consistent through time. Therefore, we apply an offset correction of –1.5‰ to all ST wheel measurements in our further analysis.

Table 1 Mean $\Delta^{14}C$ values and standard errors for air cylinder secondary standards. *Rafter All* and *Difference Rafter–INSTAAR* include results from EA and ST, with –1.5‰ bias correction applied to the ST measurements.

Cylinder name	Rafter EA only	Rafter ST only (no bias correction)	Difference Rafter ST-EA (no bias correction)	RAFTER All (ST bias corrected)	INSTAAR	Difference Rafter–INSTAAR (ST bias corrected)
NWT3	44.9 ± 0.3	45.3 ± 0.7	0.4 ± 0.8	44.7 ± 0.2	43.3 ± 0.2	1.4 ± 0.3
NWT4	–30.2 ± 0.3	–28.3 ± 0.7	1.9 ± 0.8	–30.1 ± 0.2	–31.4 ± 0.2	1.3 ± 0.3
BHDamb2013	30.7 ± 0.3	33.0 ± 0.6	2.3 ± 0.7	30.9 ± 0.2		
BHDspike2013	–72.8 ± 0.3	–70.1 ± 0.6	2.7 ± 0.7	–72.4 ± 0.2		
NIWAair	65.0 ± 0.3	66.0 ± 0.3	1.0 ± 0.4	64.7 ± 0.2		
BHD_1Nov00	92.6 ± 0.4	93.4 ± 0.3	0.8 ± 0.5	92.1 ± 0.2		
BHD_11Mar01	85.3 ± 0.4	86.5 ± 0.5	1.2 ± 0.6	85.2 ± 0.3		
BHD_10May03	80.4 ± 0.5	82.3 ± 0.8	1.9 ± 0.9	80.5 ± 0.4		
BHD_18Nov10	50.0 ± 0.6	53.6 ± 0.5	3.6 ± 0.8	51.4 ± 0.4		
Mean difference			1.5 ± 0.2			1.4 ± 0.2

Since NWT3 and NWT4 are extracted exclusively at INSTAAR, any measurement differences between Rafter and INSTAAR results must be due to graphitization and AMS measurement, including any differences resulting from OxI preparation, measurement, and normalization. INSTAAR samples are graphitized at INSTAAR and measured at the University of California Irvine AMS facility, and as with the Rafter measurements, the fractionation correction is made using on-line ^{13}C measurement, so fractionation correction method is an unlikely cause of interlaboratory differences. There is an apparent scale offset between samples prepared in the two laboratories: Rafter results (after applying our bias correction for the ST wheels) are higher by 1.4 ± 0.2‰ (Table 1). As we have already noted differences in mean $\Delta^{14}C$ depending on OxI combustion method at Rafter, the observed interlaboratory difference may well be due to differences in OxI preparation methods at INSTAAR and Rafter. However, we cannot rule out other explanations such as differences in graphitization or AMS measurement procedures.

In the FARI intercomparison of atmospheric materials (Miller et al. 2013), INSTAAR was within 0.5‰ of the interlaboratory weighted mean for one material and within 2‰ for a second material

(Rafter is participating in ongoing rounds of the FARI intercomparison, but results have not yet been compiled). Some NWT3 and NWT4 aliquots prepared by INSTAAR are also graphitized and measured at the Lawrence Livermore Center for AMS (CAMS), and showed CAMS results lower than INSTAAR by 0.2 and 0.5‰, respectively. The Rafter–INSTAAR difference of 1.4 ± 0.2‰ is similar to the spread of the FARI interlaboratory differences, but is significantly larger than the WMO recommendation for interlaboratory differences of not more than 0.5‰ (Tans and Zellweger 2014). Reducing or eliminating interlaboratory offsets for high-precision atmospheric $^{14}CO_2$ studies remains a crucial goal, since large-scale measurement coverage will require merging of data from different laboratories. For now, we recommend applying the mean offset of –1.4 ± 0.2‰ to Rafter results when using data sets reported by INSTAAR and Rafter together.

Based on analysis, we now always use OxI prepared by EA as the primary standard for air materials, and also apply this to other sample types that are not combusted (e.g. carbonates). For organic samples that require combustion, we use EA combustion wherever possible, but ST combustion is still occasionally necessary (e.g. sediments), and we match the combustion method for the samples and the primary standard.

AMS Statistical Uncertainty

The AMS statistical uncertainty (σ_{AMS}) is determined within our CALAMS software, which is described in detail by Zondervan et al. (2015) and summarized here. The dominant control on σ_{AMS} is the target Poisson uncertainty, determined from the total number of ^{14}C atoms counted from the target (i.e. $1/\sqrt{n}$), which may vary by target depending on factors including the length of measurement time, the beam current, and the mass of graphite available. σ_{AMS} also includes the Poisson uncertainty due to counting statistics on the primary standard targets, which is ~0.4‰ for our wheels with eight OxI targets generating at least 5,200,000 ^{14}C counts. We use CALAMS to test if individual exposures and/or targets show deviations that warrant outlier rejection or adjustment of σ_{AMS}. Exposure-to-exposure variability is examined for each target separately, as well as exposure-to-exposure variability for the entire wheel data set, allowing us to examine whether the excess scatter is related to a single target or to the overall AMS performance. Chi-square tests are used to evaluate the consistency of the data set by examining the scatter of the mean values and their assigned uncertainties. When the chi-square right-tail probability is less than 25% for the full data set, or less than 2.5% for a single target, we increase σ_{AMS} to account for the excess variability. This is necessary only occasionally. CALAMS also checks for consistency among all OxI primary standard targets, allowing for the addition of a "system error" associated with scatter among the OxI targets. We do not include any uncertainty associated with this in our atmospheric wheels, as we determine in the following section that this OxI scatter is an artifact of the OxI combustion process.

For our samples with 650,000 ^{14}C counts, σ_{AMS} will typically be 1.3‰ in $\Delta^{14}C$, including 1.2‰ target and 0.4 ‰ primary standard Poisson statistics added in quadrature. In cases where we extended the measurement time to obtain up to 1,200,000 ^{14}C counts, σ_{AMS} is smaller than 1.0‰. Targets containing less than 0.3 mg C are often exhausted before 650,000 counts are obtained, and consequently have larger σ_{AMS}.

Within-Wheel Repeatability

We evaluate the within-wheel repeatability, which we define as the scatter of targets of the same material measured within the same measurement wheel and normalized to the same set of primary standard targets. This allows us to determine the contribution of uncertainty due to variability in CO_2 extraction, graphitization, and individual target AMS performance. The within-wheel repeatability is evaluated by first calculating the within-wheel standardized residual r_{wwi} (Figure 2), the deviation

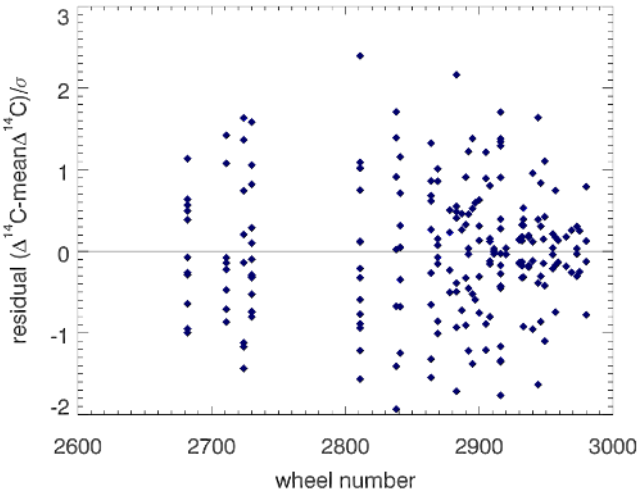

Figure 2 Within-wheel repeatability for all air standard materials measured at Rafter since 2011. Standardized residuals are calculated as the deviation of the target $\Delta^{14}C$ from the mean $\Delta^{14}C$ of all targets of that material measured within the same wheel, divided by σ_{AMS} (σ_{bw} is not included in within-wheel repeatability).

of each target from the mean $\Delta^{14}C$ of all targets of that material within the same wheel, divided by the initially assigned uncertainty σ_{AMS} for that target *i*, such that

$$r_{wwi} = \frac{(\Delta^{14}C_i - \overline{\Delta^{14}C_{ww}})}{\sigma_{AMSi}} \quad (1)$$

If the assigned uncertainties are appropriate and random, then these residuals should be normally distributed with unity standard deviation around a zero mean. We use the reduced chi-square statistic (χ^2_γ) to evaluate this, determined from all within-wheel residuals, such that

$$\chi^2_\gamma = \frac{\sum_i r^2_{wwi}}{\gamma} \quad (2)$$

In this case, the degrees of freedom (γ) is the total number of targets minus the number of groups, and here each group is a set of targets of the same material within the same wheel. A χ^2_γ of unity indicates that the assigned σ_{AMS} values are sufficient to explain the scatter of the mean values. That is, the residuals show a Gaussian distribution with a standard deviation of 1. A value of χ^2_γ greater than 1 indicates that the assigned uncertainties are not sufficient to explain the scatter and implies that the assigned uncertainties should be increased. A χ^2_γ of <1 suggests that the assigned σ_{AMS} may be too large.

We also determine the pooled standard deviation (σ_p) for the within-wheel analysis, which represents the standard deviation of all measurements, and accounting for the fact that each group has a different mean, such that

$$\sigma^2_p = \frac{\sum(N_j - 1)\sigma^2_j}{\gamma} \quad (3)$$

where σ_j is the standard deviation of the means for each material/wheel group, N_j is the number of targets in the group, and γ is the same as for χ^2_γ.

Within-wheel residuals are shown in Figure 2 and summary statistics are given in Table 2. In the within-wheel analysis, there is no EA–ST offset, since each wheel is normalized to its own primary standard targets. For the within-wheel repeatability analysis, we find that χ^2_γ is 1.0 for the air materials examined together, and σ_p is 1.3‰, the same as the typical reported σ_{AMS}. There is some variation in χ^2_γ when individual materials or groups of materials are considered separately, commensurate with the expected spread of χ^2_ν and σ_p values for small data sets, but there is no indication that particular material types exhibit more scatter than others (Table 2). We find the same result for the small subset of samples for which σ_{AMS} is 0.9–1.2‰, and for which any additional scatter would be more apparent. Thus, for air materials no additional within-wheel uncertainty is required to explain the within-wheel repeatability. Any variability due to extraction, graphitization, and measurement is apparently small enough to be undetectable at our typical σ_{AMS} of 1.3‰. The within-wheel variability contribution could be up to 0.3‰ and remain undetectable at our σ_{AMS} precision.

Table 2 Within-wheel repeatability σ_p (in ‰) and χ^2_γ values for air materials and OxI primary standard measured at Rafter since 2011. Note that the total number of targets is lower for this analysis than for the long-term repeatability, as only those where two or more targets of the same material were measured in the same wheel can be included.

Within-wheel repeatability	*n*	σ_p	χ^2_γ
Air materials	217	1.3	1.0
NWT cylinders	42	1.2	1.0
NWT3	23	1.3	1.1
NWT4	19	1.2	0.9
BHD cylinders	65	1.2	0.7
BHDamb2013	39	1.0	0.5
BHDspike2013	26	1.4	1.0
BHD split samples	124	1.4	1.1
NIWAair	46	1.5	1.1
BHD_1Nov00	33	1.4	1.3
BHD_11Mar01	13	0.9	0.6
BHD_10May03	8	1.4	1.3
BHD_18Nov10	10	1.7	1.7
OxI	350	1.6	1.6
OxI EA	224	1.5	1.6
OxI ST	126	1.8	1.6

The within-wheel repeatability analysis gives χ^2_γ of 1.6 for OxI and pooled standard deviation of 1.6‰, quite different from that of the air materials, and indicating that an additional within-wheel uncertainty of 1.0‰ is needed to explain the scatter of the OxI values. As the graphitization and AMS measurement are performed in the same way for both OxI and the air materials, the additional scatter must be due to the combustion and purification steps that are required when preparing the solid OxI material, steps which the air materials do not experience. Similar observations have been made by other authors (Graven et al. 2007; Lehman et al. 2011). OxI within-wheel repeatability appears to be slightly better for EA combustion than for ST combustion wheels.

Long-Term Repeatability

The long-term repeatability is determined from the long-term scatter of targets of the same material

across all wheels. The long-term repeatability captures the overall variability due to sample preparation and measurement, including any between-wheel variability in AMS performance. We do not consider variability due to sample collection in this analysis. It is calculated from the long-term residual r_{bwi}, the deviation of each target from the long-term mean $\Delta^{14}C$ of all targets of that material across all wheels (Table 3, Figure 3). In this calculation, we have applied the –1.5‰ bias correction to the ST wheel results.

Table 3 Long-term repeatability σ_p, χ^2_γ, and σ_{bw} for air materials measured at Rafter since 2011. A bias correction of –1.5‰ has been applied to all sealed-tube (ST) measurements.

	EA				ST				All (ST bias corrected)			
Long-term repeatability	*n*	σ_p	χ^2_γ	σ_{bw}	*n*	σ_p	χ^2_γ	σ_{bw}	*n*	σ_p	χ^2_γ	σ_{bw}
Air materials	155	1.7	1.7	1.1	110	2.0	2.0	1.3	265	1.8	1.8	1.2
NWT cylinders	52	1.6	1.5		17	2.4	2.2		69	1.8	1.7	
NWT3	28	1.6	1.6		9	2.8	3.2		37	1.9	2.1	
NWT4	24	1.6	1.3		8	1.8	1.2		32	1.6	1.2	
BHD cylinders	51	1.8	1.9		21	2.1	1.9		72	1.9	1.9	
BHDamb2013	30	1.9	2.1		11	2.1	1.9		41	2.0	2.0	
BHDspike2013	21	1.7	1.5		10	2.1	1.9		39	1.9	1.8	
BHD splits	52	1.7	1.8		72	1.9	1.9		124	1.8	1.9	
NIWAair	21	1.4	1.3		28	2.3	2.5		49	1.9	2.0	
BHD_1Nov00	10	1.8	1.8		25	1.7	1.7		35	1.7	1.8	
BHD_11Mar01	9	1.5	1.9		7	1.2	0.9		16	1.3	1.4	
BHD_10May03	8	2.6	3.1		3	1.0	0.5		11	2.3	2.3	
BHD_18Nov10	4	1.4	1.3		9	1.8	1.8		13	1.9	2.1	

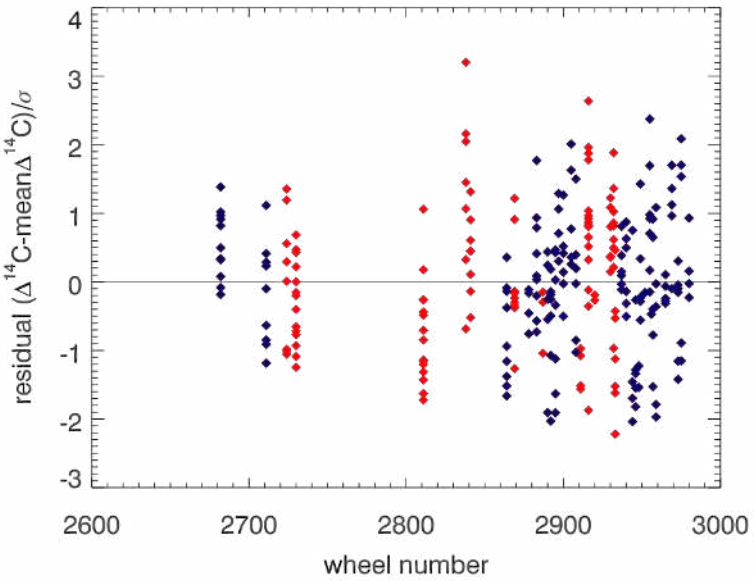

Figure 3 Long-term repeatability for all air standard materials measured at Rafter since 2011. Standardized residuals are calculated as the deviation of the target $\Delta^{14}C$ from the long-term mean $\Delta^{14}C$ for all targets of that material, divided by σ_{tot} for that target. σ_{bw} of 0.12% is included in σ_{tot} (see text). Color indicates the combustion method by which the accompanying primary standards (OxI) were prepared, either sealed tube (ST, red) or elemental analyzer (EA, blue). A bias correction of –1.5‰ has been applied to all ST wheel results.

We calculate

$$r_{bwi} = \frac{(\Delta^{14}C_i - \overline{\Delta^{14}C_{bw}})}{\sigma_{AMSi}} \quad (4)$$

and

$$\chi^2_\gamma = \frac{\sum_i r^2_{bwi}}{\gamma} \quad (5)$$

Here, γ is the total number of samples minus the number of different materials. χ^2_γ for repeatability on various subsets of the data is shown in Table 2. In all cases, σ_{AMS} is insufficient to explain the long-term repeatability. We estimate the between wheel uncertainty (σ_{bw}) required to explain the additional scatter from

$$\sigma_{bw} = \sigma_{AMS} \sqrt{(\chi^2_\gamma - 1)} \quad (6)$$

We also calculate σ_p for the long-term repeatability (Equation 3). σ_{bw} can also be thought of as the additional uncertainty (added in quadrature to σ_{AMS}) required to explain σ_p.

We find that for the air materials, we require σ_{bw} of 1.2‰. We add this σ_{bw} value in quadrature to σ_{AMS} for each individual result to obtain σ_{tot}, resulting in a typical σ_{tot} for a sample measured to 650,000 counts of 1.8‰. This is, as expected, comparable to σ_p. For wheels where the OxI was prepared by EA combustion, χ^2_γ is slightly improved at 1.7, requiring σ_{bw} of 1.1‰, and resulting in a typical σ_{tot} of 1.7‰ for EA wheels (Table 2). Wheels using ST-combusted OxI have more long-term variability, with χ^2_γ of 2.0, requiring σ_{bw} of 1.3‰.

We apply our overall between-wheel uncertainty σ_{bw} of 1.2‰ to all air measurements, regardless of primary standard method, giving A final σ_{tot} of 1.8‰ for samples measured to 650,000 ^{14}C counts. We anticipate that in the near future as more data are gathered, we will reduce σ_{bw} to reflect the improvement we identified in changing to consistently using EA OxI. In many cases, our ^{14}C measurements are used to calculate fossil-fuel CO_2 (CO_2ff), using the difference between a background and observed sample collected at approximately the same time. Wherever possible, we include the background and observed samples in the same measurement wheel, so that the between-wheel variability does not have to be considered when calculating CO_2ff. In this case, it is acceptable to use σ_{AMS} rather than σ_{tot}. In publicly available data sets, we always report the larger σ_{tot} uncertainty, to ensure that the smaller σ_{AMS} is not misapplied by later users of the data. An explanation of when σ_{bw} may be removed is included in the metadata.

CONCLUSIONS

High-precision ^{14}C measurements of air materials at Rafter give a long-term repeatability of 1.8‰, apparently improving to 1.7‰ in 2014 measurements. This includes the AMS statistical uncertainty of ~1.3‰ and a between-wheel uncertainty determined from long-term repeatability of 1.2‰. This additional between-wheel uncertainty appears to be due to scatter in the OxI primary standard material. Rafter measurements appear to be 1.4‰ higher than those made at INSTAAR, likely due to differences in preparation of the OxI primary standard.

Overall precision could be improved by extended AMS counting. In the cases where we measured to >1,000,000 ^{14}C counts and σ_{AMS} of 0.9‰, we determined that σ_{bw} of 1.2‰ is still appropriate, resulting in σ_{tot} of 1.5‰. However, this requires a near doubling of per-target measurement time from

our current high-precision configuration with 650,000 counts. In our current AMS configuration, samples of 0.5 mg C have often sputtered through the graphite material before 1,000,000 counts can be acquired. Further, the slight improvement in uncertainty needs to be weighed against the reduced sample throughput possible with extended counting times. In most cases, we think that measurement of two authentic samples to 1.8‰ is more powerful than a single measurement taken to 1.5‰.

Our results suggest that gains in long-term repeatability can more practically be made by improvements in the primary standard material preparation. We are working towards using a large bulb of OxI for air and other noncombusted materials, split under equilibrium conditions into aliquots for graphitization. We anticipate that this should reduce the scatter among OxI targets. We plan to monitor for long-term drift in the large bulb by regularly including single combustions (EA or ST) of OxI. An alternative method for improved repeatability of air materials would be to move to using cylinder standards for routine standardization within wheels, tying these cylinders to OxI regularly over time (Turnbull et al. 2013). A barrier to this is the finite lifetime of air cylinders and possible cylinder drift through time. Coordination among several or all ^{14}C laboratories measuring air materials would be ideal if this method is to be adopted.

ACKNOWLEDGMENTS

This research has been funded by the New Zealand Government (GNS-540GCT21). Our thanks to the excellent technical staff at Rafter Radiocarbon and INSTAAR for their ongoing careful laboratory work that supports these results: Kelly Lyons, Helen Zhang, Christine Prior, Cathy Ginnane, Chad Wolak, Patrick Cappa, and Steven Morgan.

REFERENCES

Baisden WT, Prior CA, Chambers D, Canessa S, Phillips A, Bertrand C, Zondervan A, Turnbull JC. 2013. Radiocarbon sample preparation and data flow at Rafter: accommodating enhanced throughput and precision. *Nuclear Instruments and Methods in Physics Research B* 294:194–8.

Conway TJ, Lang PM, Masarie KA. 2011. Atmospheric carbon dioxide dry air mole fractions from the NOAA/ESRL Carbon Cycle Global Cooperative Network, 1968–2010, version 2011-06-21. Path: ftp://ftp.cmdl.noaa.gov/ccg/co2/flask/event/.

Currie KI, Brailsford G, Nichol S, Gomez A, Sparks R, Lassey KR, Riedel K. 2009. Tropospheric $^{14}CO_2$ at Wellington, New Zealand: the world's longest record. *Biogeochemistry* 104(1–3):5–22.

Donahue DJ, Linick T, Jull AJT. 1990. Isotope-ratio and background corrections for accelerator mass spectrometry radiocarbon measurements. *Radiocarbon* 32(2):135–42.

Graven HD, Guilderson TP, Keeling RF. 2007. Methods for high-precision ^{14}C AMS measurements of atmospheric CO_2 at LLNL. *Radiocarbon* 49(2):349–56.

Graven HD, Gruber N, Key R, Khatiwala S, Giraud X. 2012. Changing controls on oceanic radiocarbon: new insights on shallow-to-deep ocean exchange and anthropogenic CO_2 uptake. *Journal of Geophysical Research* 117(C10):C10005.

Lehman SJ, Miller JB, Turnbull JC, Southon JR, Tans PP, Sweeney C. 2011. $^{14}CO_2$ measurements in the NOAA/ESRL Global Co-operative Sampling Network: an update on measurements and data quality. In: Brand WA, editor. *Expert Group Recommendations from the 15th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases and Related Tracer Measurement Techniques*. Jena: World Meterological Organisation. p 315–7.

Lehman SJ, Miller JB, Wolak C, Southon JR, Tans PP, Montzka SA, Sweeney C, Andrews AE, LaFranchi BW, Guilderson TP. 2013. Allocation of terrestrial carbon sources using $^{14}CO_2$: methods, measurement, and modelling. *Radiocarbon* 55(2–3):1484–95.

Levin I, Kromer B. 1997. Twenty years of atmospheric $^{14}CO_2$ observations at Schauinsland station, Germany. *Radiocarbon* 39(2):205–18.

Levin I, Kromer B, Schmidt M, Sartorius H. 2003. A novel approach for independent budgeting of fossil fuel CO_2 over Europe by $^{14}CO_2$ observations. *Geophysical Research Letters* 30(23):2194.

Lowe DC, Judd W. 1987. Graphite target preparation for radiocarbon dating by accelerator mass spectrometry. *Nuclear Instruments and Methods in Physics Research B* 28(1):113–6.

Manning MR, Lowe DC, Melhuish WH, Sparks RJ, Wallace G, Brenninkmeijer CAM, McGill RC. 1990. The use of radiocarbon measurements in atmospheric sciences. *Radiocarbon* 32(1):37–58.

Miller JB, Lehman S, Wolak C, Turnbull J, Dunn G, Graven H, Keeling R, Meijer HAJ, Aerts-Bijma AT, Palstra SWL, Smith AM, Allison C, Southon J, Xu X, Nakazawa T, Aoki S, Nakamura T, Guilderson T,

LaFranchi B, Mukai H, Terao Y, Uchida M, Kondo M. 2013. Initial results of an intercomparison of AMS-based atmospheric $^{14}CO_2$ measurements. *Radiocarbon* 55(2–3):1475–83.

Rafter TA, Fergusson G. 1959. Atmospheric radiocarbon as a tracer in geophysical circulation problems. In: *United Nations Peaceful Uses of Atomic Energy*. London: Pergamon Press. p 526–32.

Reimer PJ, Brown TA, Reimer RW. 2004. Discussion: reporting and calibration of post-bomb ^{14}C data. *Radiocarbon* 46(3):1299–304.

Stuiver M, Polach HA. 1977. Discussion: reporting of ^{14}C data. *Radiocarbon* 19(3):355–63.

Tans PP, Zellweger C. 2014. *17th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques (GGMT-2013)*. World Meteorological Organization Global Atmosphere Watch.

Turnbull JC, Lehman SJ, Miller JB, Sparks RJ, Southon JR, Tans PP. 2007. A new high precision $^{14}CO_2$ time series for North American continental air. *Journal of Geophysical Research* 112:D11310.

Turnbull JC, Guenther D, Karion A, Sweeney C, Anderson E, Andrews AE, Kofler J, Miles NL, Newberger T, Richardson SJ. 2012. An integrated flask sample collection system for greenhouse gas measurements. *Atmospheric Measurement Techniques* 5:2321–7.

Turnbull JC, Graven HD, Miller JB, Lehman SJ. 2013. Atmospheric radiocarbon workshop report. *Radiocarbon* 55(2–3):1470–4.

Turnbull JC, Lehman SJ, Morgan S, Wolak C. 2010. A new automated extraction system for ^{14}C measurement in atmospheric CO_2. *Radiocarbon* 52(3):1261–9.

Van Der Laan S, Karstens U, Neubert REM, Van Der Laan-Luijkx IT, Meijer HAJ. 2010. Observation-based estimates of fossil fuel-derived CO_2 emissions in the Netherlands using $\Delta^{14}C$, CO and ^{222}Radon. *Tellus B* 62(5):389–402.

Zondervan A, Hauser T, Kaiser J, Kitchen R, Turnbull JC, West JG. 2015. XCAMS: the compact ^{14}C accelerator mass spectrometer extended for ^{10}Be and ^{26}Al at GNS Science, New Zealand. *Nuclear Instruments and Methods*. doi:10.1016/j.nimb.2015.03.013.