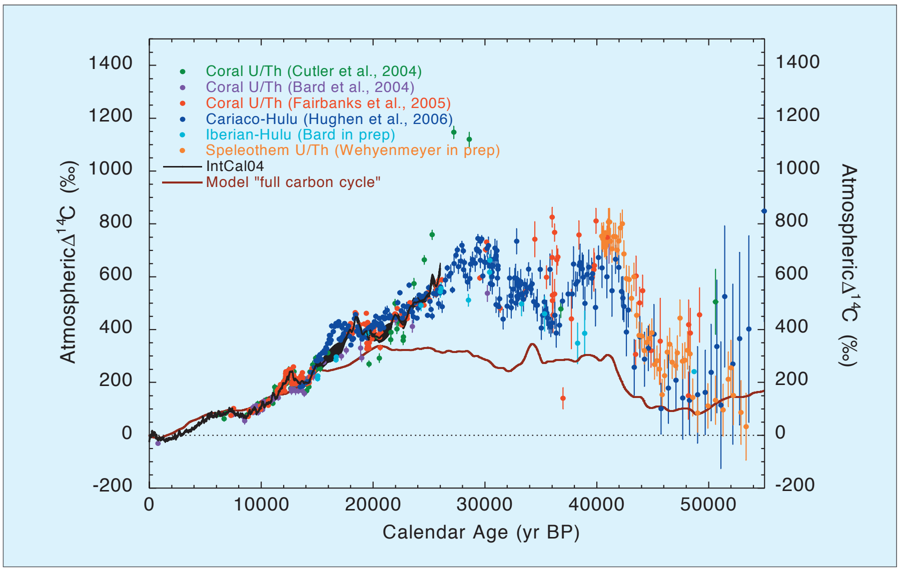

Figure 2: ^{14}C activity ($\Delta^{14}C$) data for the past 50,000 years. IntCal04 is shown back to its limit of 26 kyr. Marine coral and sediment ^{14}C data have been corrected with a constant reservoir age, and speleothem data have been corrected with a constant dead carbon fraction. The paleo-$\Delta^{14}C$ observations are plotted compared to a carbon cycle box model simulation representing fixed preindustrial boundary conditions and changing ^{14}C production ("full carbon cycle"). Error bars show 1-σ $\Delta^{14}C$ uncertainty.

$\Delta^{14}C$ response of ~200‰, encompassing most of the Glacial age data. Reducing flux to shallow sediment reservoirs is required to match the highest observed $\Delta^{14}C$ values. According to the model, however, the prescribed change in surface-to-deep-ocean exchange would produce a doubling of the surface-to-deep-ocean $\Delta^{14}C$ difference. Observations do provide some evidence of decreased Glacial $\Delta^{14}C$ in the deep western and eastern North Atlantic, as well as deep eastern equatorial and southwest Pacific (for review see Hughen et al., 2006). However, such a large change in Glacial deep ocean $\Delta^{14}C$ has not been observed in the western equatorial Pacific (e.g. Broecker et al., 2004). It is important to note that most of the paleo-ocean $\Delta^{14}C$ reconstructions correspond to the period around the Last Glacial Maximum (~21 kyr BP), an interval when the simulated $\Delta^{14}C$ response to production rate changes alone is close to the observations (especially if reasonable production rate uncertainties are considered). Another serious issue is that reconstructed rates of $\Delta^{14}C$ change at the beginning of the last deglaciation, ~17 kyr, are too large to be explained by changes in production rate alone and require a substantial dilution of ^{14}C atoms in the atmosphere by a more depleted reservoir. Reconstructions of transient deglacial $\Delta^{14}C$ changes in the intermediate depth western and deep eastern North Atlantic are consistent with a major reorganization of deep ocean circulation at that time, probably involving increased ventilation of a previously isolated deep water mass of southern or Pacific origin (e.g. Adkins et al., 2002). These model simulations can place constraints on the magnitude of deep ocean $\Delta^{14}C$ anomalies required to explain the surface marine record. In addition, the model data make quantitative predictions of the increase in surface marine reservoir age during the Glacial period. Unfortunately, however, observations of Glacial reservoir variability from low-latitude sites are rare. More sophisticated model simulations with increased spatial resolution would help identify the patterns of increase in reservoir age according to latitude and ocean basin. Regardless, it is clear that high-quality observations are needed from each of the three principal carbon reservoirs—atmosphere, surface and deep ocean—in order to constrain changes in both deep ocean $\Delta^{14}C$ and surface marine reservoir age, and to understand the history of radiocarbon and global carbon cycle changes.

References

Broecker, W., Barker, S., Clark, E., Hajdas, I., Bonani, G. and Lowell, S., 2004: Ventilation of the glacial deep Pacific Ocean, *Science*, **306**: 1169-1172.

Hughen, K.A., Southon, J.A., Lehman, S.J., Bertrand, C.J.H. and Turnbull, J., 2006: Marine-Derived ^{14}C Calibration and Activity Record for the past 50,000 Years Updated from the Cariaco Basin, *Quaternary Science Reviews*, in press.

Laj, C., Kissel, C. and Beer, J., 2004: High resolution global paleointensity stack since 75 kyr (GLOPIS-75) calibrated to absolute values, in: *Timescales of the Paleomagnetic Field, Channell, J., Kent, D., Lowrie, W. and Meert, J. (Eds.)*, American Geophysical Union, Washington, D.C., 255-265.

Masarik, J. and Beer, J., 1999: Simulation of particle fluxes and cosmogenic nuclide production in the Earth's atmosphere, *Journal of Geophysical Research*, **104**: 12,099-12,112.

Muscheler, R., Beer, J., Kubik, P.W. and Synal, H.A., 2005: Geomagnetic field intensity during the last 60,000 years based on ^{10}Be and ^{36}Cl from the Summit ice cores and ^{14}C, *Quaternary Science Reviews*, **24**, 1849-1860.

For full references please consult:
www.pages-igbp.org/products/newsletters/ref2006_3.html

Assuring measurement quality: The international ^{14}C laboratory inter-comparison program

E. M. Scott

Department of Statistics, University of Glasgow, UK; marian@stats.gla.ac.uk

Introduction

In order to achieve reliable, precise and accurate ^{14}C age measurements, laboratories routinely undertake both formal and informal quality assurance programs. Such programs may involve the repeated and routine measurement of an internal standard (such as a bulk cellulose sample), the results of which enable the laboratory to evaluate their reliability and precision. They may also routinely have access to known-age material against which to assess their accuracy. Beyond this, however, many laboratories regularly participate in inter-laboratory comparisons to provide independent checks on laboratory performance.

Reference material for ^{14}C calibration

High-quality ^{14}C measurements also require traceability to international standards whose ^{14}C-activities are known exactly by independent means, and also to reference materials whose activities are estimated and typically accompanied by associated uncertainty statements. Within the ^{14}C community, there has been an increasing realization of the need for adequate reference materials and a resultant development of both internal and external quality assurance (QA) procedures. Routinely, ^{14}C laboratories make use of a number of standards and reference materials whose activities are known or are estimated from large numbers of measurements made by many laboratories (e.g. NIST OxI, OxII, IAEA C1-C8). More recent ^{14}C inter-comparisons have also created a further series of natural reference materials (Scott, 2003, Scott et al, in prep).

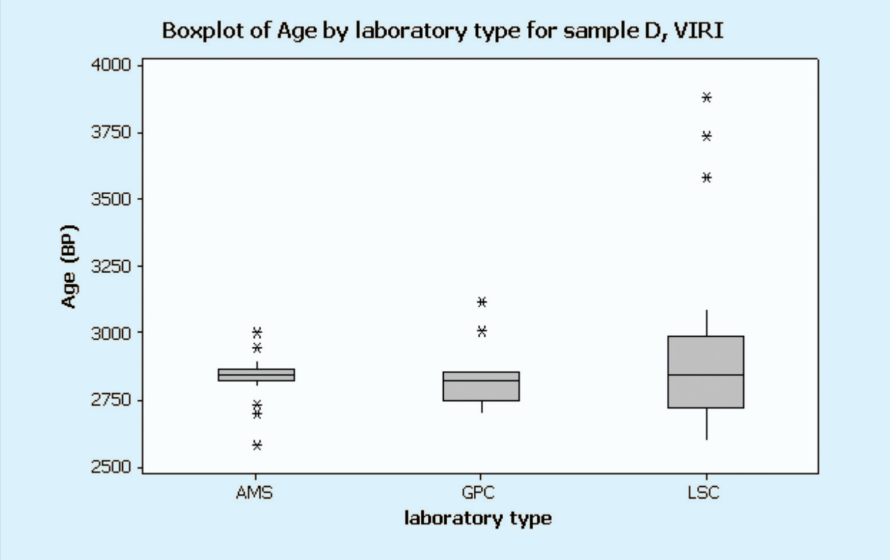

*Figure 1: Boxplot showing the distribution of results for a charred grain sample (sample D, VIRI) for the three laboratory types (32 AMS, 31 LSC and 10 GPC). The boxplot represents the median, lower and upper quartile and minimum and maximum. An * identifies an outlier. The lower quartile is the value below which 25% of the results lie.*

History of inter-comparisons

Since the early days of applied ^{14}C measurement, it has been common practice for laboratories to exchange samples in attempts to improve and sustain analytical confidence. With time, this practice has given way to a succession of more formal inter-comparison exercises. Within the ^{14}C community in the past 20 years, there have been a number of very extensive inter-laboratory trials. These comparisons have varied widely in terms of sample type and preparation, with their primary goal the investigation of the comparability of results produced under possibly quite different laboratory protocols. Regular comparisons have reassured users that the laboratories are striving to ensure highest quality results, while at the same time allowing the laboratories to identify any systematic offsets and additional sources of variation. Thus, participation in a laboratory inter-comparison has become an important part of a formal QA program.

Here we summarize some of the findings from the two most recent ^{14}C inter-comparison exercises (FIRI (Fourth) and VIRI (Fifth), Scott 2003 and Scott et al, in prep.). VIRI is ongoing at this time but continues the tradition of TIRI (Third) and FIRI, operating as an independent check on laboratory procedures. It is a 4-year project, with the first phase already completed. Phase 2, using bone samples, is due to be reported by the end of 2006. Further stages will include samples of peat, wood and shell with a range of ages. VIRI, like the TIRI and FIRI inter-comparisons, is a ^{14}C community project, with samples provided by participants and a substantial laboratory participation rate of over 75%.

What questions have been asked and answered by these inter-comparisons?

- Comparability of measurements from different laboratories

One of the main questions that such inter-comparisons are used to answer concerns how comparable the results are among laboratories, especially where some use different procedures and techniques. For the ^{14}C dating community, this historically concerned the comparability of results from the accelerator mass spectrometric (AMS) and radiometric (liquid scintillation (LSC) or gas proportional counting (GPC)) laboratories.

From FIRI, overall and on average, no evidence of significant differences in the results between AMS, GPC and LSC laboratories was found. In the first phase of VIRI, a similar preliminary conclusion was drawn (Scott et al, in prep.). Figure 1 below shows the distribution of results for sample D (charred grain) in VIRI.

- Variation

Clearly the results among laboratories do vary, but an inter-comparison exercise can assess the degree of variation and also which factors might explain such variation (aside from simply random fluctuations). One aspect of variation concerns outliers, or extreme observations. In FIRI, roughly 10% of the total results were identified as outliers (which is around twice as frequent as would be expected). An outlier is an observation which is either too young or too old, defined statistically as those values that are greater than 3 inter-quartile ranges from the nearest of either the lower or upper quartiles. The distribution of outliers was not homogeneous across the laboratories, with the majority of outliers coming from around only 14% of the laboratories.

Can we identify any reasons for the variation in results? Modern standard and background material used were studied but no evidence was found that these factors made a significant contribution to the overall variation. The type of modern standard used, however, did seem associated with the outlier distribution.

- Accuracy

Accuracy is concerned with the 'correctness of the result' and refers to the deviation (difference) of the measured value from the 'true' value. Ideally with known-age samples, this can be independently estimated and a small number of known-age samples have been included (typically dendrochronologically dated wood). However, more commonly, we must assume that we

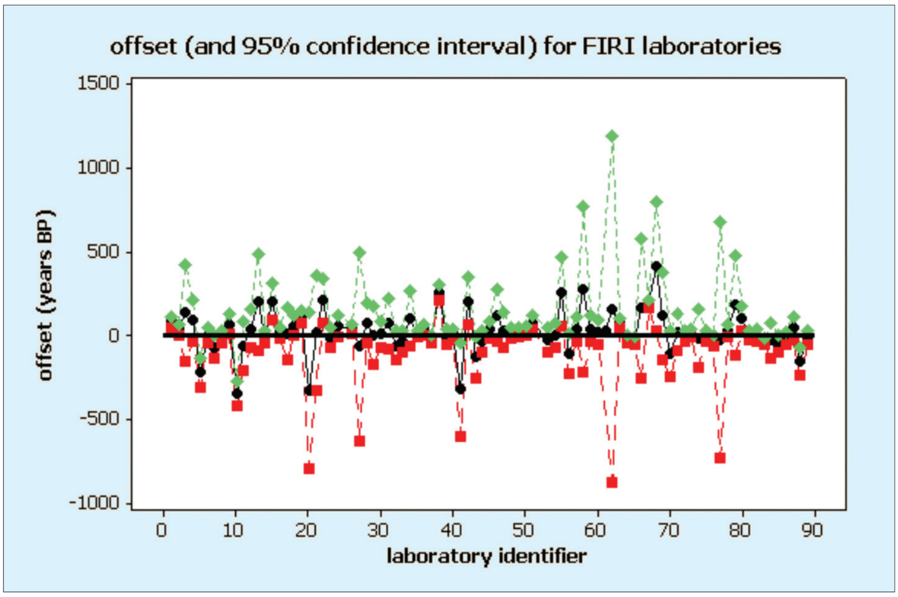

Figure 2: Offset (and 95% confidence interval) for dendrochronologically dated samples (3200-3239 BC, 3299 - 3257 BC and 313-294 BC) showing the upper (green diamonds) and lower (red squares) limits of the 95% confidence interval, and the point estimates of the laboratory offsets (black circles).

can define (through calculation) what the 'true' ^{14}C age will be (the consensus value) and then we can estimate for each laboratory, whether there is a constant offset (or a bias) from this consensus. The current program of inter-laboratory comparisons has afforded an opportunity for laboratories to assess their accuracy. In each intercomparison, the consensus values for the unknown age samples was calculated and reported. Figure 2, , shows the offset (and 95% confidence interval) for individual laboratories based on the dendrochronologically dated samples included in FIRI. The sample dendro-ages were 3200-3239 BC, 3299 - 3257 BC and 313-294 BC.

Conclusions

Analyses of results from FIRI and phase 1 of VIRI support the fact that radiocarbon laboratories are generally accurate and precise. The results from FIRI are significant in that they show broad agreement between measurements made in different laboratories on a wide range of materials, and they also demonstrate no statistically significant difference between measurements made by radiometric or AMS techniques. As a result of the inter-comparison program, an extensive suite of natural reference materials (such as wood, carbonate, etc) spanning the applied ^{14}C timescale has been created by the ^{14}C dating community. These can now be used by ^{14}C labs to improve their dating accuracy and are thus of great benefit to the users of ^{14}C dates.

Acknowledgements

FIRI and VIRI are supported by EC FPIV programme and English Heritage. The assistance of sample providers is gratefully acknowledged, in addition to the support of all participating labs.

References

Scott, E.M., 2003: The Third International Radiocarbon Intercomparison (TIRI) and The Fourth International Radiocarbon Intercomparison (FIRI), *Radiocarbon*, **45**(2):135-328.

Scott, E.M., Cook, G.T., Naysmith, P., Bryant, C. and O'Donnell, D., A report on phase 1 of the 5th International Radiocarbon inter-comparison (VIRI), *Radiocarbon*, in prep.

IntCal and the future of radiocarbon calibration

P. J. Reimer[1], E. Bard[2], C. Buck[3], T. P. Guilderson[4,5], A. Hogg[6], K. Hughen[7], B. Kromer[8], R. Reimer[1], J. Southon[9], C. S. M. Turney[10], J. van der Plicht[11,12], C. E. Weyhenmeyer[13] and C. B. Ramsey[14]

[1]CHRONO Centre for Climate, the Environment, and Chronology, Queen's University Belfast, UK; p.j.reimer@qub.ac.uk

[2]CEREGE, UMR-6635, Aix-en-Provence, France; [3]Department of Probability and Statistics, University of Sheffield, UK; [4]Center for Accelerator Mass Spectrometry L-397, Lawrence Livermore National Laboratory, USA; [5]Ocean Sciences Department, University of California, Santa Cruz, USA; [6]Radiocarbon Dating Laboratory, University of Waikato, Hamilton, New Zealand; [7]Department of Marine Chemistry & Geochemistry, Woods Hole Oceanographic Institution, USA; [8]Heidelberger Akademie der Wissenschaften, Heidelberg, Germany; [9]Department of Earth System Science, University of California, Irvine, USA; [10]GeoQuEST Research Centre, School of Earth and Environmental Sciences, University of Wollongong, Australia; [11]Centrum voor Isotopen Onderzoek, Rijksuniversiteit Groningen, Groningen, Netherlands; [12]Faculty of Archaeology, Leiden University, Netherlands; [13]Department of Earth Sciences, Syracuse University, USA; [14]Oxford Radiocarbon Accelerator Unit, UK

Background

In addition to being crucial to many archaeological studies, radiocarbon ages form the chronological basis for many Holocene and late Pleistocene paleoclimatic studies and paleoenvironmental reconstructions. The basic radiocarbon age calculation assumption of constant atmospheric ^{14}C content is not valid, however, due to solar- and geomagnetic-induced changes in production rate and ocean circulation changes. In order to compare radiocarbon chronologies with those derived from other means, such as ice core or U/Th dated sequences, it is necessary to calibrate against measurements of "known" age samples.

Calibration curves were originally based only on ^{14}C measurements of known-age tree-rings and a calibration curve for Holocene marine samples was constructed using the atmospheric data as input into a simple ocean-atmosphere box diffusion model. More recently, marine records, U-Th dated corals and foraminifera from varve-counted sediments of Cariaco Basin, combined with reservoir corrections, provide high-resolution atmospheric calibration beyond the range of the tree-ring record. The ocean-atmosphere box diffusion box model, however, is used for Holocene marine calibration where marine calibration data are generally not available with sufficient resolution and precision. After a disastrous start of multiple independent "calibration" data sets that yielded disparate calendar ages and confusion among users, radiocarbon calibration curves have been constructed by small groups utilizing the best available and reproducible data sets, and ratified by the attendees at the International Radiocarbon Conferences. As the types of records available for calibration have diversified and the pros and cons of each realized, a larger, more formalized group has become a necessity. In 2002, the IntCal Working Group was created and has since met at a series of workshops funded by the Leverhulme Trust. The IntCal group has produced estimates of the calibration curves for the main carbon reservoirs: the Northern Hemisphere atmosphere (IntCal04), the Southern Hemisphere atmosphere (SHCal04), and the marine environment (Marine04). SHCal04 was the first ratified calibration curve for the Southern Hemisphere.

Most of the early calibration curves were constructed from a simple weighted

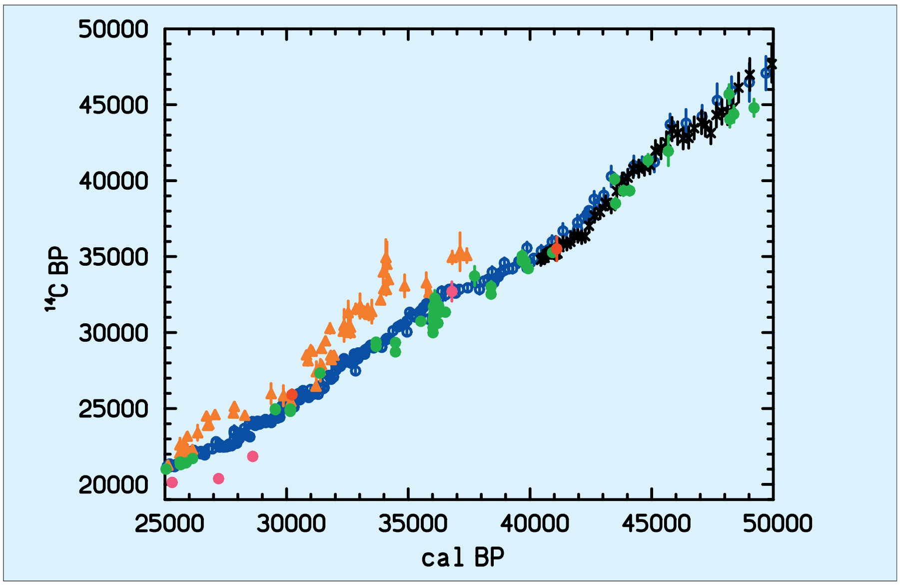

Figure 1: Selected ^{14}C data sets >26 cal kyr BP with calendar timescales based on U/Th dating, varve counts or through correlation to U/Th-dated speleothems. The marine and speleothem records are corrected with a constant reservoir or dead carbon fraction offset as specified in the original publication. Uncertainty is shown in the ^{14}C ages only as one standard deviation. Corals are given as solid circles: red (Bard et al., 2004); pink (Cutler et al., 2004); green (Fairbanks et al., 2005). The Cariaco Basin foraminifera data with the timescale from the correlation to the Hulu Cave is given as open blue circles (Hughen et al., 2006), Lake Suigetsu macrofossils as solid orange triangles (Kitagawa and van der Plicht, 2000), and Arabian Sea speleothem as black crosses (Weyhenmeyer et al., 2003).