PRELIMINARY REPORT OF THE FIRST WORKSHOP OF THE INTCAL04 RADIOCARBON CALIBRATION/COMPARISON WORKING GROUP

Paula J Reimer 1 • Konrad A Hughen 2 • Thomas P Guilderson 1 • Gerry McCormac 3 • Mike G L Baillie 3 • Edouard Bard 4 • Phillip Barratt 3 • J Warren Beck 5 • Caitlin E Buck 6 • Paul E Damon 7 • Michael Friedrich 8 • Bernd Kromer 9 • Christopher Bronk Ramsey 10 • Ron W Reimer 3 • Sabine Remmele 8 • John R Southon 11 • Minze Stuiver 12 • Johannes van der Plicht 13

ABSTRACT. The first meeting of the IntCal04 working group took place at Queen’s University Belfast from April 15 to 17, 2002. The participants are listed as co-authors of this report. The meeting considered criteria for the acceptance of data into the next official calibration dataset, the importance of including reliable estimates of uncertainty in both the radiocarbon ages and the cal ages, and potential methods for combining datasets. This preliminary report summarizes the criteria that were discussed, but does not yet give specific recommendations for inclusion or exclusion of individual datasets.

INTRODUCTION

Calibrated radiocarbon ages are the basis for comparison between many records in paleoclimatological, geophysical, and archaeological studies. It is, therefore, necessary to have a standard 14C calibration dataset for intercomparisons to be valid. The 14C community has recognized this need and hence has adopted an international standard for calibration, most recently IntCal98 (Stuiver et al. 1998a), which was ratified at the 16th International Radiocarbon Conference in Groningen (van der Plicht and Mook 1998). The importance of the quality of the dataset used for calibration cannot be overstated. As new data become available, careful consideration must be given to whether or not they should be included in the official calibration dataset. A balance must be maintained between accepting only true “calibration” datasets (i.e. those where the age on the cal axis is known very precisely such as dendrochronological dated tree rings and U/Th dated pristine corals) and incorporating all available datasets. A conservative approach has been followed in the past to avoid inclusion of datasets with potential errors in the cal ages or other underlying assumptions. While this means that some information about the details of the calibration dataset may be unavailable, the conservative approach assures that general trends are reliable. Therefore, we are advocating such a conservative approach for IntCal in the future.

1 Center for Accelerator Mass Spectrometry L-397, Lawrence Livermore National Laboratory, Livermore, California 94550, USA. Corresponding author. Email: pjreimer@llnl.gov.
2 Woods Hole Oceanographic Institution, Department of Marine Chemistry & Geochemistry, Woods Hole, Massachusetts 02543, USA.
3 School of Archaeology and Palaeoecology, The Queen’s University of Belfast, Belfast BT7 1NN, Northern Ireland.
4 CEREGE, UMR-6635, Europole de l’Arbois BP80, 13545 Aix-en-Provence cdx 4, France.
5 NSF–Arizona AMS Facility, Department of Physics, University of Arizona, Tucson, Arizona 85721, USA.
6 Department of Probability and Statistics, University of Sheffield, Sheffield, S3 7RH, England.
7 Department of Geosciences, PO Box 210077, University of Arizona, Tucson, Arizona 85721, USA.
8 Universität Hohenheim, Institut für Botanik-210, D-70593 Stuttgart, Germany.
9 Heidelberger Akademie der Wissenschaften, Im Neuenheimer Feld 229, D-69120 Heidelberg, Germany.
11 Department of Earth System Science, University of California–Irvine, Irvine, California 92697, USA.
12 Quaternary Isotope Lab, University of Washington, Seattle, Washington 98195, USA.
13 Centrum voor Isotopen Onderzoek, Rijksuniversiteit Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands.
CRITERIA FOR ACCEPTANCE FOR CALIBRATION AND COMPARISON DATASETS

In general, for acceptance into the IntCal compilation, an individual dataset must pass certain criteria that depend on the type of record and measurements made. In all cases, uncertainties in both the 14C age and the cal age must be quantified. A review of the datasets included in the IntCal98 calibration dataset and discussion of some problems with various “comparison” datasets are given by van der Plicht (2000). While some otherwise good datasets may not meet all criteria for inclusion into the calibration dataset, they may still be valuable for comparisons and for checking the magnitude of atmospheric excursions and/or marine reservoir corrections.

Datasets with large analytical errors in either the cal age or the 14C age will only add noise and will be excluded. New data must be as good as extant IntCal98 for the time period in question. We would like input on what analytical errors would be considered acceptable for older samples where we know it is difficult to obtain highly precise and accurate measurements. It is not likely that we will develop a calibration record with ±30 yr 14C errors to 50,000 yr!

Tree-Ring Criteria

Tree rings should be dendrochronologically dated and cross-checked where possible. In some cases, trees with a known felling age may be used by ring counting, if rings are well developed and single annual ring production in the species is known to be robust. For instance, the Douglas-firs (Pseudotsuga menziesii) from the Pacific Northwest were ring counted back to AD 1320 (Stuiver 1982). Uncertainty in single ring cal ages for dendrochronologically dated wood is on the order of 1 yr for highly replicated and crosschecked chronologies. An additional small uncertainty in cal age exists for 14C dates on multiple yr blocks of wood, because of an unequal amount of carbon in each ring. In extreme cases, this could result in an uncertainty of several 14C yr only. Floating tree-chronologies have been used in the past with a wiggle-matched connection to the dendrochronological secured chronology (Stuiver et al. 1998a). While this approach could be problematic, we note that the German pine wiggle-match ~10,000 cal BP used in IntCal98 was only 8 yr different from the dendrochronological result as detailed in a later section.

Coral Criteria

X-ray diffraction measurements should be performed and should show ≤1% calcite as a check against secondary calcite. The [U] of fossil corals should be within the range of living/modern samples taking into account differences between coral species and the natural variability expected from sea surface temperature (SST) changes. In order to increase our confidence that the corals have not been altered by diagenesis, δ234U_initial of fossil corals should be within ±5‰ of the accepted modern seawater value. This criterion is based on current understanding that δ234U in seawater is constant over the last ~30,000 yr. Accuracy and precision should be checked by measuring and reporting the δ234U of recent corals and/or seawater (Delanghe et al. 2002). An alternative demonstration of accuracy and precision is to report δ234U of an international standard with δ234U similar to natural uranium (e.g. SRM-U960 or –U010) (Bard et al. 1998; Edwards et al. 1993). External replicates of different parts of the same coral sample should be measured as often as possible to derive an estimate of external reproducibility, which should be reported with the results. This demonstrates the overall reproducibility of U-Th ages and is also a necessary requirement suggesting closed system behavior. Interlaboratory comparisons of standards and coral samples are to be encouraged. New datasets should be encouraged to include protactinium measurement in particular for samples that have been subaerially exposed in the past. While it is impractical to measure protactinium for all previous coral calibration datasets, if discrepancies arise, we would encourage reanalysis of a sub-set of the original
samples. AMS-14C coral samples need to be physically separated and subjected to partial dissolution on the order of 30–50% with the degree of leaching necessary to be determined by experiment. Background correction must be applied and background errors included in the error analysis. External replicates should be done wherever possible to derive a more accurate external reproducibility.

The error bar of U-Th ages should be on the order or less than that obtained by 14C measurement on the same sample. Consequently, samples counted by alpha spectrometry do not have sufficient precision to be included in IntCal. Very precise measurements do not necessarily mean accurate—we expect that there will be instances where there will be results that we cannot explain even if they meet our geochemical criteria. In part, this is a consequence of the subtleties of diagenesis and it will more likely impact those samples that have been subaerially exposed or have been recovered from fast uplifting regions.

A site-specific reservoir correction should be estimated with a “reasonable” error, if at all possible, although this may prove to be untenable for time periods of rapidly changing atmospheric 14C. There is a preference for future records to be developed from oceanographically “simple” regions to minimize reservoir age uncertainty. Coral datasets should have a well-described and published stratigraphy. The ages should be distributed in a logical manner along the stratigraphic sequence, taking into account rates of sea-level changes, vertical movements (subsidence or uplift), the paleobathymetry of the coral species and the topography of the reef substratum.

Non-Coral Carbonates Criteria

Records of non-coral carbonates, such as flowstones, stalagmites, and aragonite deposits, have additional uncertainties in both the U/Th ages and 14C ages. This is the result of possible incorporation of detrital Th and old or dead carbon from carbonate dissolution processes (e.g. water-rock interaction) or other sources of old carbon in lacustrine/marine environments. The initial conditions of the carbon budget need to be confirmed by independent environmental indicators (e.g. δ13C). The U/Th age model should be verified by layer counting or 14C measurements of terrestrial macrofossils, where possible. Appropriate corrections to measured values need to be applied. It is important to evaluate these conditions for various growth/deposition periods as the relative contribution of detrital Th and dead/old carbon may significantly change through time as a result of changes in the climate regime (e.g. ocean circulation, precipitation, temperature, vegetation). As with corals, partial dissolution of carbonates may be necessary to remove contamination, but the degree of leaching required should be determined for each particular material.

Non-Dendro Layer Counting Criteria (Terrestrial and Marine)

Multiple-core chronologies are critical in order to confirm that no sections are missing (e.g. from core-breaks, erosional scour, etc.). A “multiple proxy” approach in counting should be used. Discussion of errors should include uncertainty about the stability of the system producing layers through climatic changes, the likelihood of unrecognized depositional hiatuses, and any ambiguities in layer interpretation that are not resolvable by even “perfect” physical archive recovery (e.g. could lakes have frozen over during colder climates?). It is not sufficient to report only counting statistics. For both laminated sediments and ice cores, it is important that counting be replicated, whenever possible, by experienced individuals in independent laboratories (Hicks et al. 1991). Confirmation of the counting should also be made with independent chronologies and/or tie-points. Ideally, individual varves would be correlated across cores (Ralska-Jasiewiczowa et al. 1998). We acknowledge that there will be some sets of data that might not have the possibility of physical replication via multiple cores and in this instance independent chronologies and/or tie-points are increasingly important. An
obvious example of this would be geochemical identification of precisely dated ash horizons (Gronvold et al. 1995).

There should be a minimal level of acceptable scatter in macrofossil 14C ages due to reworking or delayed transport of old organic material, although this should be evaluated “piecewise”. This means that a particular varved record could still be useful even if it includes a section with relatively high scatter. The type of macrofossils used will undoubtedly vary, but should be reported for each sample, as it influences the interpretation of how long it takes the macrofossil from “creation” to deposition in the sediment. Where possible, multiple macrofossils should be analyzed. For marine sequences, the reservoir correction and the uncertainty in reservoir changes over time must be estimated.

Reconstructed Time Scales

Records with a calendar timescale based on correlation with signals in annually counted ice cores, or sediments with assumed constant 14C time/depth relationship and absolute time markers have been used to reconstruct “comparison” curves. Many of these are marine records, and criteria for minimum acceptable boundary conditions are needed. Hydrographically complex locations, such as inside the polar front or in upwelling regions where we expect large variations in local reservoir age, are not ideal. A site should have simple 14C systematics not requiring complex circulation or ventilation dynamics (i.e. to explain disparate benthic/planktonic data). Defendable and reasonable estimates of minimal bioturbation should be made—for example using ash shard counts. As the chronology of many of these records is based on correlation of one or more proxies to ice core properties (GISP2 and GRIP), it is critical to identify the physical mechanism relating the proxy to the ice core property, as well as the timescale of the linking mechanism. In addition, the errors in the ice core timescales must be considered. We encourage the continued investigation of differences between GISP2 and GRIP (and eventually NGRIP) timescales, as well as identification of radiometrically dated tie-points (e.g. Laschamps geomagnetic excursion; Dansgaard-Oeschger climate events).

Geomagnetic and Cosmogenic Isotope Based Models

Models based on paleomagnetic intensity or 10Be and the global carbon cycle are useful correlation tools and can be important indicators of the causes of 14C variations. However, at present, our understanding of cosmogenic production and the global carbon cycle, as well as 10Be transport, deposition, and post-depositional processes, is not sufficient to make such records suitable for 14C calibration.

PROJECTIONS FOR INTCAL04

Refinements to IntCal98 for the period 0–12 ka BP were readily agreed upon by workshop participants and are presented below. These changes should result in only minor differences from IntCal98 through the Holocene. Between 12 and 15 ka BP, there is good agreement between datasets, several of which meet the criteria laid out in the earlier section, therefore, a “calibration” curve can be given with reasonable certainty, although there are some details which remain to be worked out. From 15 to 25 ka BP there are important differences between some of the existing datasets with only coral data truly meeting the established criteria. However, some additional detailed coral data should be available for this time period (Cutler et al. forthcoming). These new data, together with an alternative statistical technique to the spline fit of coral data used for IntCal98, will hopefully remove the very large uncertainty at around 15 ka BP, which results in the feature known as the “pig-in-the-python” (Figure 1).

Although the working group is not yet ready to make specific recommendations for IntCal04 beyond 25 ka BP, there was much discussion about the need to provide some guidelines, if not calibration datasets, for the entire 14C timescale.
The Holocene part of the $^{14}$C calibration is based on several millennia-long tree-ring chronologies providing an annual, absolute time frame, which was rigorously tested by internal replication of many overlapping sections. Whenever possible, they were crosschecked with independently established chronologies of adjacent regions. The German and Irish oak chronologies were cross-dated until back into the 3rd millennium BC (Pilcher et al. 1984), and the German oak chronologies from the Main River, built independently in the Göttingen and Hohenheim tree-ring laboratories, cross-date back to 9147 cal BP (Spurk et al. 1998). The North American trees that form part of the $^{14}$C calibration curve prior to AD 1320 were cross-dated with either the *Sequoiadendron* master chronology or with a Pacific Northwest Douglas-fir chronology (Stuiver 1982).

Before 9147 cal BP, we have to rely on single oak (back to 10,430 cal BP) and pine chronologies of the Hohenheim laboratory, which are not replicated externally. However, for the earliest oak the internal replication is high, and the trees cross-date with high statistical significance. Compared to the state entered into IntCal98, the German pine chronology has undergone several corrections:

1. The $^{14}$C wiggle-match of the floating pine to the absolute oak chronology is superseded by a true dendro-synchronization of the 2 species, leading to a 8-yr shift of the pine with respect to the $^{14}$C match (which had a ±20 yr confidence interval) (Friedrich et al. 1999).
2. The earliest centuries (prior to 11,200 cal BP) in the pine chronology, marked as tentative in Spurk et al. (1998), suffered severely from missing rings and a weak dendro-match. Using ample new finds, this part of the chronology is now securely cross-dated and sufficiently replicated. The tentative dendro-link as documented in Spurk et al. (1998) in the interval 11,200 to 11,370 cal BP is replaced by a statistically reliable match. The older part of the chronology therefore was shifted 70 yr to older ages. The former start of the German pine at 9922 cal BC (11,871 cal BP) (Spurk et al. 1998) is now shifted to 9992 cal BC (11,941 cal BP).
3. Including additional new finds the German pine is prolonged into the Younger Dryas. It now starts at 11,993 cal BP. A 20-tree pine chronology from Lake Neufchatel extends it further back to 12,057 cal BP.

Thus, while most of the 14C calibration data as based on the German pine is affected only marginally by the corrections (+8 yr), in the first century of the Preboreal, and in the final five centuries of Younger Dryas the dendro-dates are shifted by 70 yr to older ages, compared to IntCal98. This interval happens to fall on a gently sloping 14C age plateau (at 10,000–10,150 14C BP); hence the length of the plateau and the range of calibrated ages increase accordingly.

Because of periodic narrow rings caused by cockchafer beetles some German oak samples were excluded from IntCal98. Analysis of these tree-rings, with an understanding of the response of trees to the cockchafer damage, may allow some of these measurements to be re-instated in the chronology. Wood will be made available for measurement for any remaining gaps in the tree-ring 14C record resulting from previous shifts.

Recent measurements of Belfast Irish oak over the past 1000 yr (Hogg et al. 2002) will be included to increase precision during this time period. Comparisons of these and other measurements with the 1986 and 1993 Irish oak datasets resulted in the acceptance of the 1986 datasets over the 1993 datasets. A Southern Hemisphere specific dendro-calibrated record will be included using the measurements of the New Zealand trees over the past 1000 yr (Hogg et al. 2002) and other Southern Hemisphere datasets (McCormac et al. 2002). An offset of $41 \pm 14$ 14C yr BP from the IntCal98 dataset from AD 950 to 1850 will be used to extend the record through the Holocene. The offset will be recalculated for use with the new IntCal dataset. While this Southern Hemisphere 14C offset may not be representative of the entire Holocene, we note that measurements for the early Holocene are of the same order of magnitude (Kromer et al. 1998).

The Arizona Bristlecone pine chronology (Linick et al. 1986), which was not included in IntCal98, was also discussed. Wiggle-matching the bristlecone pine dataset to IntCal98 suggests that there is not a problem with the dendrochronology but there is an average offset of approximately 35 yr in the 14C ages with the bristlecone pine older. Part of this offset may be due to laboratory differences as the offset between 15 decades of bristlecone pine measured in both Seattle and Arizona is $25 \pm 8$ yr with Arizona measurements being older. However, comparisons made in several laboratories of bristlecone pine to German oak and Irish oak find the bristlecone pine consistently older. Furthermore, measurements in Arizona of single yr rings of sequoia from the Sequoia National Forest averaged as decades from AD 1065 to 1145 resulted in a difference of only $-1.0 \pm 2.1$ 14C yr BP compared to Seattle measurements of Douglas-fir from the Pacific Northwest. Therefore, while a regional or growing season difference during some time periods at least seems possible, more work is needed to resolve the issue. The treatment of offsets between datasets, and the inclusion of the Arizona Bristlecone pine measurements, will be discussed further at the next working group meeting. Additional tree-ring records will be considered for inclusion if available by the time of the next workshop.

The marine model used to produce the IntCal98 marine dataset from 0 to 8800 cal BP (Stuiver et al. 1999b) will be further tested by checking the model response to the nuclear-testing 14C spike and to centennial-scale variations. The variability in the marine reservoir observed in recent records will be considered to better represent the uncertainty in the marine dataset for this period. Coral data (Bard et al. 1998; Burr et al. 1998; Edwards et al. 1993) and the Cariaco Basin varve dataset will provide the marine calibration from 8.8 to 12 ka BP. The high resolution Cariaco Basin dataset (Hughen et al. 2000) will replace the older varve dataset in IntCal98. The cal time scale will be optimized using appropriate coral tie-points within the error of the varve counting. As in previous versions of the
marine calibration dataset, the time-dependent difference, $\Delta R(t)$, between the reservoir age of the region where samples are derived and the “global” calibration dataset, will need to be considered when calibrating marine samples.

12–15 ka BP

The coral datasets, including any new records meeting the IntCal criteria, and the Cariaco Basin varve dataset will provide the marine calibration for this period. The varved records from Lake Suigetsu in Japan (Kitagawa and van der Plicht 2000) and Lake Gościąż in Poland (Goslar et al. 2000) agree well with the “site-specific” reservoir corrected corals (T Goslar, personal communication 2000) and the new Cariaco Basin dataset but are not as high resolution as the Cariaco Basin and exhibit larger variability. It is anticipated that the atmospheric dataset for this portion of the curve will be based on the marine dataset with “site-specific” reservoir corrections and estimated uncertainties. The extent to which the reservoir corrections may vary with time will be estimated by comparison to lake varve macrofossil data, floating tree-rings, and other suitable paleorecords.

15–25 ka BP

Coral data provide the only secure calibration for this portion of the timescale. Although there is reasonable agreement between parts of the Bahamas speleothems (Beck et al. 2001) and the Lake Suigetsu varve record for this time period, the $^{14}$C ages differ by up to 1200 yr at 20 ka BP suggesting that there is an error in either the Lake Suigetsu varve ages or the assumption of a constant dead carbon fraction in the Bahamas speleothem or a combination of both. At present, other available paleorecords either have cal age uncertainty that is too large or difficult to quantify, or require untested assumptions about the stability of the system producing the records. If other records or additional support for current records become available, they will be reconsidered at the next IntCal workshop.

Beyond 25 ka BP

Great strides have recently been made towards generating $^{14}$C datasets beyond 25 ka BP with U/Th measured cal ages (Beck et al. 2001; Kitagawa and van der Plicht 2000; Schramp et al. 2000; Yokoyama et al. 2000) and with matches to the ice core timescales (van Kreveld et al. 2000; Voelker et al. 2000). However, these records are not sufficiently coherent that we can recommend a single calibration for this time period. Individually, many of these records reveal large and intriguing structure; however, differences of more than 5000 yr are observed amongst these records in some time frames. The choice of a particular dataset could result in vastly different “calibrated” age ranges (Bard 2001). The sources of these discrepancies are not currently understood, but may be due to some combination of artifacts relating to varve counting, $^{230}$Th dating, uncertainties in the GISP2 chronology or correlation with GISP2 $\delta^{18}$O, unaccounted for variability in reservoir correction or dead carbon fraction, or undetected overprinting from secondary alteration or authigenic mineral growth. Combining these various records together into a composite calibration would result in such a large error envelope as to make the calibration useless. Still, leaving the $^{14}$C user-community to use “ad hoc” mixtures of $^{14}$C records without proper consideration of potential errors in timescale or reservoir corrections is inadvisable, and could result in substantial confusion and misinterpretation. No decision was made at the Belfast IntCal04 working group meeting regarding how to resolve this problem, though it will be again addressed at the next workshop at Woods Hole Oceanographic Institution (WHOI). At the very least, we can advise the user community of the danger of ad-hoc calibration without proper consideration of all sources of uncertainty. Alternatively, we could attempt to provide a set of comparison curves that includes all quantifiable uncertainties in both the cal age and the $^{14}$C age for these records, though these parameters may be difficult to assess for some cases.
Ideally, we would provide tools for the statistical estimation of cal age ranges given the types of information available to us. In any case, we recommend that no endorsement of a calibration in this time range be made at the present.

Post-Bomb Data

It was suggested that IntCal provide a compilation of post-nuclear testing atmospheric records. However, as the records are latitude dependent and many are not from clean air sites, it may be impractical to combine datasets. It would however be useful to provide access to all known post-bomb records in a single location with appropriate errors included. While some of this is already done at the CDIAC (Carbon Dioxide Information Analysis Center) as well as the NGDC “Paleo” database, we will attempt to provide an updated list of where post-bomb records can be found. We encourage individual investigators to digitally archive their respective datasets at the appropriate site, and this information should be provided in any original manuscripts.

Methods of Combining Data

In quantifying and reporting the error on the cal scale for the first time, we will need to reconsider the methods by which the data are combined to form the calibration curve. Thus, as well as conventional statistical methods of summarization, we will investigate the use of stochastic models for building calibration curves.

PRESENTATION OF RESULTS

It is difficult for the working group to predict what uses will be made of the data once they are released. As a result, and to provide as much flexibility as possible, the group agreed that it would be important to make the raw calibration data available in a suitable form for use in spectral analysis and curve modeling as well as to provide a recommended curve for immediate calibration purposes. It is anticipated that the IntCal04 calibration dataset will be presented for ratification at the 18th International Radiocarbon Conference in Wellington, New Zealand in 2003 prior to publication in early 2004.

ACKNOWLEDGMENTS

We wish to thank the Leverhulme Trust for support for this workshop and the compilation of the anticipated IntCal dataset. We also want to thank Prof Gerry McCormac of the Radiocarbon Lab at Queen’s University Belfast for the generous hospitality that contributed to the success of the meeting. A portion of this work was performed under the auspices of the US Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

REFERENCES


