Radiocarbon, Vol 56, Nr 2, 2014, p 567–572 DOI: 10.2458/56.14447

A NEW DATABASE PROGRAM INSTALLED AT THE SUERC RADIOCARBON LABORATORY

Brian G Tripney[1] • Philip Naysmith • Gordon T Cook

Scottish Universities Environmental Research Centre, East Kilbride G75 0QF, Scotland.

ABSTRACT. The SUERC Radiocarbon Dating Laboratory has recently replaced its spreadsheet-based record keeping with a new database program, custom designed to help laboratory staff manage the high throughput of nearly 5000 cathodes in the past year. The system can accept data from a variety of sources in addition to manual entry; experimental results can be uploaded from spreadsheets, while integration with graphitization lines means that graphite yields are automatically recorded. The system is able to pass ^{14}C results directly to OxCal v 4 for calibration, with the resulting plots incorporated into the dating certificates issued to submitters. There are also benefits to submitters, with electronic sample submission both eliminating transcription errors and speeding up the logging-in process, which keeps turnaround times down. For bone samples, data on collagen yields are now stored electronically and are more readily obtainable from the laboratory. The new SUERC Radiocarbon Dating Laboratory database will make a significant contribution to maintaining the high quality of results produced by the laboratory, aiding staff in tracking sample progress, and monitoring quality assurance (QA) samples going through the laboratory, eliminating transcription errors, and making communication easier between laboratory staff and sample submitters.

INTRODUCTION

Originally set up at Glasgow University in 1967, the Radiocarbon Dating Laboratory moved to East Kilbride in 1986 and has since been based at SUERC. Liquid scintillation counting was the focus for the laboratory until 1996 when it began to produce graphite targets for measurement at the University of Arizona Accelerator Mass Spectrometry Laboratory. Following the installation of a NEC 5MV tandem AMS at SUERC in 2003, target measurement was moved in-house and in 2007 a 250kV single-stage AMS (SSAMS) was added to provide additional ^{14}C measurement capability.

As of 1998, there has been an electronic record of all samples processed and the results obtained, stored in annually produced spreadsheets. However, growing demand for ^{14}C dating services has led to an increased throughput of samples with the laboratory now preparing approximately 3300 unknowns and 1700 quality assurance (QA) samples for AMS measurement each year. This high volume of samples proved increasingly cumbersome to manage using a spreadsheet, particularly when attempting to retrieve details for samples received some months apart or producing dating certificates through a mail merge. These drawbacks and the emerging need for concurrent access to the data by staff have led the laboratory to replace the spreadsheet-based system with the new database system described below.

SAMPLE SUBMISSION

The system allows sample details (e.g. material type, species, sample ID) to be automatically read in from a submission spreadsheet and prefilled into the relevant sections of the *New Submission* window. This helps to eliminate transcription errors, which means that, assuming the information provided by the submitter is accurate, the correct sample details will appear throughout the system and eventually on the dating certificate.

Where possible, a member of staff will match the submitter contact details to an existing record; otherwise, new entries are created in the database. At this point, the staff member will enter the recommended pretreatment scheme for each sample, set the requested turnaround time, and ensure

1. Corresponding author. Brian.Tripney@glasgow.ac.uk.

Proceedings of the Radiocarbon and Archaeology 7th International Symposium
Ghent, Belgium, April 2013 | Edited by Mark Van Strydonck, Philippe Crombé, and Guy De Mulder

that the appropriate stable isotopes are selected for analysis. The next available reference number is assigned to track the sample through the various laboratory processes. These laboratory numbers each begin with the prefix GU- to signify processing at the SUERC Radiocarbon Laboratory.

Having saved the sample details in the database, the system then prints the laboratory log sheets that will accompany each sample through the laboratory, detailing the processes to be applied and providing space for staff members to record yields at different stages in the sample preparation process, notes on pretreatment, etc. At this point, the database also generates an email to the sample submitter to confirm receipt of the samples.

This electronic submission method not only reduces the scope for data entry errors on behalf of laboratory personnel, but also joins together the previously disjointed subtasks of receiving samples, printing laboratory log sheets, and informing submitters. These improvements speed up the logging-in process and help to keep sample turnaround times down.

PROGRESS TRACKING

From this point onward, it is possible to track the status of each sample accepted for dating through the database system. The first update on sample progress is recorded in the system when pretreatment has been completed successfully. A staff member will record the completion date in the database before passing the sample on for CO_2 production and graphitization. In the case of bone, antler, and tooth samples, additional information on the collagen yield is recorded at this stage.

The progress of each sample is updated as each stage of the dating process is completed. Following the pretreatment stage, this is an automatic process when an experimental result is added to the database. The dates of completion of each stage can be seen when viewing the list of samples (Figure 1), making it easy to check how far a sample has progressed through the laboratory.

To maintain the quality of the results issued by the laboratory, a sample may be failed at any point in the process if there is reason to doubt that a reliable result can be obtained. Once a sample has been marked as a fail in the database system, it will appear on the fail management page. At this point, the system provides options to return the sample to the laboratory to try a different approach, or confirm the failure of the sample. An email containing the sample details can be generated to inform the submitter.

An overview page (Figure 2) makes use of the sample status information recorded in the database to give a summary of laboratory performance. The top row of numbers shows the total number of samples at each major stage in the process: in pretreatment, waiting for graphitization, prepared for AMS batch, being measured on AMS, and ready to report. This allows management of staff to prevent bottlenecks occurring.

Pretreatment Date	Graphite Date	Mass Spec Date	CN Date	AMS Batch Date	Date Reported	Status
07/03/2013	11/03/2013	14/03/2013	n/a	20/03/2013	28/03/2013	Reported
07/03/2013	11/03/2013	14/03/2013	18/03/2013	20/03/2013	28/03/2013	Reported
12/03/2013	18/03/2013	25/03/2013	25/03/2013	Pending		In Progress
19/03/2013						FailStableIsotopes
12/03/2013	19/03/2013	20/03/2013	25/03/2013	Pending		In Progress
12/03/2013	19/03/2013	25/03/2013	25/03/2013	Pending		In Progress
14/03/2013	19/03/2013	Pending	Pending	Pending		In Progress

Figure 1 Sample tracking in the database system

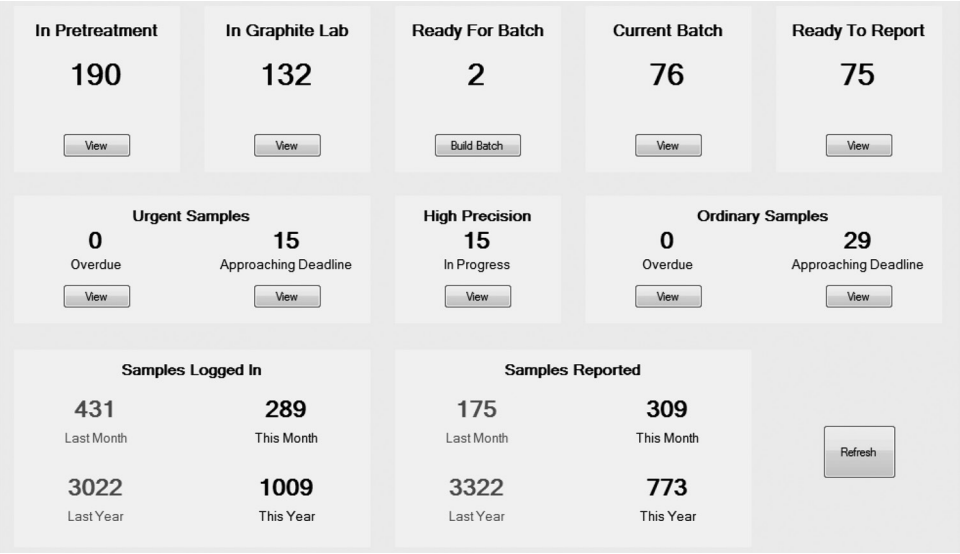

Figure 2 Laboratory overview page

The second row of the overview page highlights any samples that are approaching (or have exceeded) their deadline, so that any problems causing delay can be remedied. The final row displays the total number of unknown samples logged in and the total number reported in the current month and year, allowing the overall laboratory throughput to be seen at a glance. The totals for the previous month and year are also shown for comparison.

RECORDING EXPERIMENTAL RESULTS

The results of the processes and measurements that follow pretreatment are recorded by the system using a range of data input methods. These are described below in increasing order of automation.

As the mass spectrometers used for $\delta^{13}C$ measurement are not connected to the network, the raw measurement data from the printouts must be typed in by a member of staff. A record of each mass spectrometer calibration is stored in the database, allowing the $\delta^{13}C$ value of each sample to be calculated. The manual entry of this data is recognized as a potential for typographical errors and as such, the $\delta^{13}C$ values are recalculated "offline" and any discrepancies amended as appropriate.

For bone, antler, and tooth collagen samples, additional analysis of nitrogen isotopes is undertaken using continuous-flow isotope ratio mass spectrometry (CF-IRMS). These $\delta^{15}N$ and C/N ratio measurements are easier to enter as they are obtained in a spreadsheet format that can be read into the database in a similar manner to the submission spreadsheets.

However, the most closely integrated process is that of graphitization. The SUERC Graphite Program is linked directly to the database system and takes advantage of this during the daily setup of the graphitization lines as well as at the point of storing results. At the point of entering identification numbers of the samples to be run, the program displays the sample details as a check for the user and, once a run has been started, the list of samples currently being graphitized is available through the database program. At the completion of the graphitization run, a sheet giving sample details and plots of conversion of CO_2 to graphite is printed and the graphite yield data are transferred to the database. A staff member then decides which of the samples have successfully passed the graphitization yield criterion and records this in the database, flagging the sample as being ready to be sent to the AMS.

AMS BATCHES AND RESULTS

A typical batch of graphite targets measured on the AMS consists of 13 groups, each containing three QA samples and seven unknowns. By keeping track of sample progress, the database can assist staff in assembling batches of samples for AMS by only presenting samples that are graphitized and ready to measure for inclusion in a batch. The system displays two lists of unknown samples for the user, one for ordinary samples and another for special service samples (e.g. fast turnaround or high-precision measurement). The samples are ordered within the lists such that the samples received earliest are displayed first.

QA samples are displayed in additional lists according to type. The laboratory numbering for these QA samples is separate to the GU- coding used for unknown samples, with the prefix identifying the type of each sample. International standard oxalic acid II is denoted by prefix M (modern), while secondary reference materials use prefixes BBM (bulk barley mash) and HA (humic acid). Further prefixes are used for background material (BK), and in-house bone QA samples.

As each sample is selected for inclusion in the batch, the system automatically positions it according to the batch template, ensuring QA and unknown samples appear in the correct places. Once the batch has been filled, the paperwork to accompany the batch to the AMS is generated and the batch date is recorded against each sample.

When data are returned from the AMS, the sample ages are calculated, informed by the results received for the QA samples, and checked by staff members. Use of a standard spreadsheet ensures the final age results, including the SUERC- coding that indicates measurement at the SUERC AMS Facility, always appear in a predefined position and format to allow easy upload of the spreadsheet into the database. At this point, the system will check to ensure that all the results and details required to report each sample are available, alerting the user to any required action.

Once all the necessary information is available, the next step for the majority of samples is the calibration of the ^{14}C age to the calendar timescale using OxCal v 4 (Bronk Ramsey 2009). To achieve this, the system allows the user to select a group of samples and the appropriate atmospheric or marine curve, and passes this to OxCal using a URL-based command. The calibration plots can then be saved individually and the resulting SVG files are batch-converted to PNG format for inclusion on the dating certificates.

Samples are presented for final reporting when a converted calibration plot is detected (modern or failed samples do not require this). The dating certificates are then generated by the database system, incorporating the plot if appropriate, saved as PDF, and a hard copy sent to the printer. The content of the report is automatically customized to meet the requirements of each sample—for example, a result given as fraction modern will have a standard text block added to explain this. If so required, a set of results may alternatively be saved as a spreadsheet.

TECHNICAL DETAIL

The system is currently built around a server-based MySQL 5 database and is primarily accessed through a custom user interface installed on each user's computer. This allows for simultaneous access by multiple users. The tabbed interface splits the system functions into related groups: logging in, pretreatment, stable isotopes, quality assurance, and reporting. Additional tabs containing laboratory and database management tools are displayed to certain users only.

To provide a rich user interface with access to local printers and devices, the front end is coded in

VisualBasic.NET and is connected to the database using MySQL Connector.[2] The ability to read data in from spreadsheets is provided by GemBox Spreadsheet,[3] which allows the values of individual cells to be read or for a group of cells to be extracted to a *DataTable*. The same interface is used to save data to a spreadsheet—one cell at a time or by inserting an entire DataTable. Inkscape,[4] a vector graphics editor, is used in its command line batch-mode to convert the SVG files obtained from OxCal into a more manageable PNG format and the production of electronic versions of the dating certificates is handled through the COM interface of PDFCreator.[5]

The main database tables are shown in Figure 3. Every sample received is recorded in the *PhysicalSample* table, which holds the identifying number and type of each sample. To allow for subsampling or repeated processing of any physical sample, the *LabSample* table provides storage for a suffix to be appended to the laboratory number. It is this number (a combination of *LabPrefix*, *LabID*, and *LabSuffix*) that identifies each sample in the experimental results tables to the right of the figure.

The *MassSpecResult*, *CNResult*, and *GraphiteResult* tables allow for repeated measurement of any sample. Measurements that are not accepted or are replaced have the value of the *Result* field set to

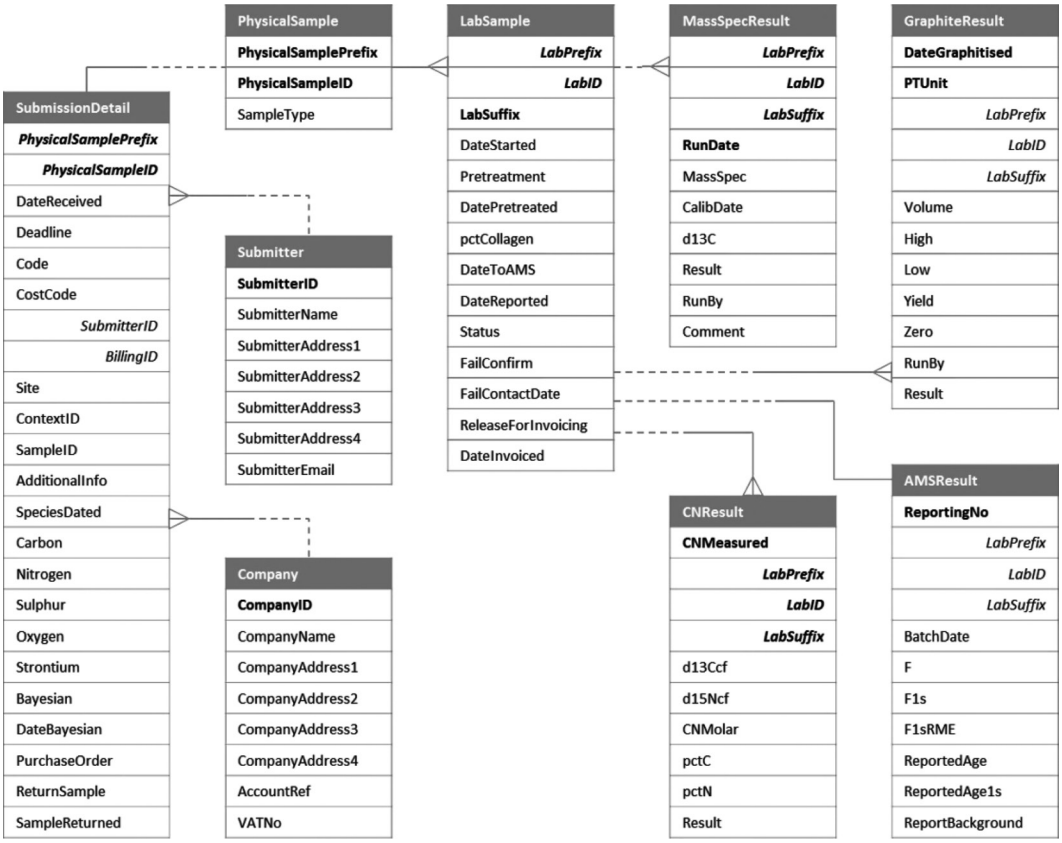

Figure 3 Relationships between main database tables

2. MySQL Connector/NET, version 6.4.3, Oracle 2011 (www.mysql.com/products/connector).
3. GemBox Spreadsheet Free, version 3.3, GemBox Software 2011 (www.gemboxsoftware.com).
4. Inkscape, version 0.48.2, Inkscape Project 2011 (www.inkscape.org).
5. PDFCreator, version 1.6.2, pdfforge 2013 (www.pdfforge.org).

"Fail," preventing its use later in the process while maintaining a record of the measurement. Only successful AMS measurements are entered into the *AMSResult* table, with the unique SUERC-number recorded as *ReportingNo*.

For submitted samples, the *SubmissionDetail* table records information on the sample, submitter, deadlines, and analyses required; use of this table is not required for QA samples. As submitters are likely to have multiple samples associated with them, it follows that their contact details should be stored separately from the sample details. Consequently, only a *SubmitterID* is stored as part of the sample information in the *SubmissionDetail* table, with the full contact details stored in the *Submitter* table. Billing details are similarly stored in the *Company* table.

In addition to the table structure shown here, a number of support tables are employed to keep track of other details relevant to the system. The next available laboratory numbers for submitted samples (GU- coding) and each QA sample type are listed in the *NextSampleNo* table, with a record of any advance payment made by a submitter stored in *Prepayment*. For convenience, *AMSBatch* summarizes batch sample totals and *CurrentGraphites* lists samples currently running on the graphitization lines.

Data integrity within tables is enforced through the use of primary keys and between tables by the use of foreign key constraints. The primary key for each table (shown in Figure 3 in bold) provides a unique identifier for each entry in a database table and normally includes the sample laboratory number. The database will not permit duplication of the primary key, so where the laboratory number may be repeated in a table it is supplemented with additional detail to form a primary key—for example, the *MassSpecResult* table uses the laboratory number and the date of measurement as its primary key. A foreign key (shown in italics in Figure 3) requires that data must exist in another table before being added to the current one, e.g. a measurement cannot be added to a results table before the sample it refers to is added to the *LabSample* table. To permit the enforcement of these foreign key constraints, the MySQL database is set up using the InnoDB table type. Additional checks on data, such as querying values outside a normal range, are enforced by the user interface before entry into the database.

CONCLUSIONS

The new SUERC Radiocarbon Database has helped to manage the increasing volume of samples processed by the laboratory, not only streamlining some time-consuming administrative parts of the process, but also aiding in the tracking of samples and recall of sample details. Developed in-house, the system is constantly evolving, with changes ranging from small tweaks to the addition of new features. Work is currently underway to develop communication with the Historic Scotland Radiocarbon Dating Recording System. Once samples are approved for funding, the sample details will be received electronically from the Historic Scotland system and passed directly to the laboratory database. By associating the samples with the unique identifiers provided by the external system, it will be possible to return the results and dating certificates through the same channel. When operational, this will considerably reduce duplication of effort on behalf of submitters, Historic Scotland, and the laboratory. Further planned work includes incorporation of the laboratory's older electronic records into the new database, providing access to samples from 1998 to the present day through a single system.

REFERENCES

Bronk Ramsey C. 2009. Bayesian analysis of radiocarbon dates. *Radiocarbon* 51(1):337–60.