Radiocarbon, Vol 56, Nr 4, 2014, p S79–S83
DOI: http://dx.doi.org/10.2458/azu_rc.56.18320

TREE-RING RESEARCH, Vol. 70(3), 2014, pp. S79–S83
DOI: http://dx.doi.org/10.3959/1536-1098-70.3.79

DATA MANAGEMENT IN DENDROARCHAEOLOGY USING TELLERVO

PETER W. BREWER

Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ 85721, USA. Corresponding author: p.brewer@ltrr.arizona.edu.

ABSTRACT

The Tellervo dendrochronological software builds upon the Tree-Ring Data Standard (TRiDaS) to provide a tool for recording and managing all manner of dendrochronological data. However, Tellervo is especially useful for dendroarchaeological research. The traditional file formats used in dendrochronology—and by association the applications that use them—have very limited and nonstandard methods for recording rich information about dendro samples and their context. Such information is especially important in dendroarchaeological research to ensure accurate conclusions are made. Tellervo is described here in the context of research carried out as part of the excavations of the Theodosian Harbor at Yenikapı, Istanbul.

Keywords: TRiDaS, Tellervo, data standards, data management, dendroarchaeology.

INTRODUCTION

The dendrochronology community has a long and rich history of using computational techniques to facilitate research. As early as 1963, mainframe computers were being routinely used at the Laboratory of Tree-Ring Research (LTRR), University of Arizona (Fritts 1963). This pioneering work by Fritts and others revolutionized the field.

In the 1960s and 1970s, the standard mechanism for transferring both software and data was through punched cards. Remarkably, the legacy of the punch card is still felt today—over half a century on—through the Tucson decadal file format, which is still commonly used to archive tree-ring width data. The format, designed in the 1970s by LTRR faculty and staff, was in many ways a digital representation of the punch cards widely used up until that point: that is, an identifier followed by 10 columns of ring-width values in an 80-column row. The Tucson format remains the most widespread format in use today including as the archiving format used by the International Tree-Ring Data Bank (ITRDB; Grissino-Mayer and Fritts 1997). The ITRDB is the world's primary repository of dendrochronology data and as such is widely regarded as the standard for storing tree-ring width and density data. The Tucson format does, however, have many limitations that present challenges for dendrochronologists, especially for those using dendrochronology for archaeological and cultural research (e.g. art history, buildings historians, provenance studies, etc).

Limitations in both computational power and storage capacity led to compromises with the design of the Tucson file format. For example, it provides just four characters for calendar year (limiting the earliest year to –999 = 1000 BC), six for series identifier, and four for each ring-width value. Note that the Tucson format uses astronomical year numbering that matches Julian/Gregorian years in the AD period, but is offset by one for years BC. The astronomical year 1 (=AD 1) is preceded by the year 0 (=1 BC), then year –1 (=2 BC) and so forth. This can cause problems for the uninitiated who assume a negative year number is directly equivalent to a BC year.

In the absence of a formal authority administering the format, software developers and users have independently introduced a number of informal alterations to work around these (as well as other) limitations. For example, the final character of the identifier is used by some as part of the calendar year value, thus providing five characters and extending support back to year –9999 (=10,000 BC). The six characters for series identifier are also often extended to eight characters, with the remaining columns shifted to the right by two. Although these and other alterations to the format are perfectly logical, support for "nonstrict" Tucson files varies between software.

The same restrictions on processing and storage capacity also resulted in very minimal provision for the storage of metadata, i.e. information describing the samples and data. The Tucson format is essentially a data-only format, with space for just three header lines for metadata. The problem is compounded by the fact that these metadata lines are only very loosely standardized. The standard layout is largely ignored within the community, resulting in these metadata fields being in essence, free-text (Brewer *et al.* 2011).

THE IMPORTANCE OF METADATA

The absence of structured metadata is a substantial problem for those working in the various fields of cultural dendrochronology, where typically a great deal of associated information is required to adequately process and interpret dendro samples. For instance, the canonical goal of a dendroarchaeology study is to determine the felling date for a tree used in a structure to enable conclusions to be drawn about the likely construction date. However, unlike trees sampled *in situ*, archaeological samples are often devoid of bark and some or all of the sapwood. A determination for the date of the final ring is therefore of limited use without information about the presence or absence of bark and/or sapwood. Likewise, the state of preservation of archaeological samples can often be poor. In some circumstances, the final rings of a sample can be counted, but are too badly damaged to measure meaningfully. Accurate recording of such information is essential when interpreting results.

Other important metadata routinely required by cultural dendrochronologists is the detailed information about where the samples were taken. Historical and archaeological sources of samples such as buildings, ships, paintings, and musical instruments may have been constructed in several phases, and/or have been repaired one or more times. In these cases, an intimate understanding and detailed records of the structure being sampled are needed. For example, the iconic Hagia Sophia in Istanbul includes tie-beams from the original 6th century Justinian construction, through to wood from the 19th century repairs overseen by Gaspare and Giuseppe Fossati (P. I. Kuniholm, personal communication, 26 March 2014). In circumstances such as this, time can be wasted and errors can be made in the interpretation of results if detailed records of each sample are not maintained.

Metadata is also essential for the reuse of dendroarchaeological data. Papers within this volume outline the pressing need for a better understanding of the interactions between humans and the environment—indeed, this is the focus of the new Center for Mediterranean Archaeology and the Environment (CMATE) at the University of Arizona. Tree rings hold great potential in this regard and have been used with great success to investigate past climates (e.g. Grudd *et al.* 2002; Friedrich *et al.* 2004; Cook *et al.* 2006; Salzer and Hughes 2007). However, in the Mediterranean region, because of the absence of long-lived tree species (such as the bristlecone pines of North America or the kauris of New Zealand) and the absence of suitable environments for the preservation of subfossil wood (such as the bogs of northwestern Europe), we are reliant on wood samples from archaeological contexts to understand climate prior to the last few hundred years.

Using archaeological samples presents many challenges for such work, not least because of the mobility of wood from the archaeological record. Although samples taken from live or *in situ* subfossil trees will record the climatic conditions from the locations where they are found, the same is not true of archaeological samples. Ships may be found hundreds of miles from where they were built, let alone from where the trees used to build them were growing. Even static objects such as buildings may have been built using materials sourced from great distances. To enable the effective use of dendroarchaeological data in environmental reconstructions, it is therefore essential that accurate and rich details about the sampled object and its provenience are maintained.

Clearly, the Tucson file format was never designed to store such detailed metadata. Over the past 3 decades, a variety of dendro file formats have been introduced to address some of these limitations, but none have been widely adopted beyond the software for which they were initially designed and have a variety of limitations (Brewer *et al.* 2011). As a result, cultural dendrochronologists typically resort to *ad hoc* procedures in their labs to keep track of this important metadata. This consists variably of record cards, paper forms, field notebooks, spreadsheets, word processing documents, maps, etc. Although such methodologies can be adequate, they are by no means efficient and are prone to user error.

TREE-RING DATA STANDARD—TRiDaS

An initiative was launched in 2006 to resolve many of the aforementioned problems by designing and implementing a new data format capable of recording all of this metadata in a standardized way. The Tree-Ring Data Standard (TRiDaS) was published following contributions from over 80 dendrochronologists, computer scientists, and users of dendrochronological data (Jansma *et al.* 2010).

TRiDaS has been implemented as an extensible markup language (XML) schema. The schema includes the ability to describe a wide range of metadata covering all the subdisciplines of dendrochronology and is aimed to be a universal format for the storage and transmission of both dendro data and metadata. A detailed discussion of the TRiDaS schema is presented by Jansma *et al.* (2010) and does not require repeating here.

A facet of TRiDaS that is of particular note, especially in dendroarchaeological research, is the implementation of the concept of "controlled vocabularies." Controlled vocabularies are used to restrict users to a list of terms for concepts such as the type of object being studied. When presented with a free-text field, users will inevitably use a wide variety of related terms and spellings when referring to the same or similar concept. For instance, a boat could be listed variably as boat, ship, barge, galley, frigate, yacht, clipper, schooner, and probably dozens of other related terms including different languages and typographical variants. By restricting users to a list of terms, this problem is minimized; however, a controlled vocabulary becomes much more useful with the inclusion of relationships between terms. For instance, frigate, clipper, schooner, and galley are specific types of ships, so if a user searches for the broad concept *ship*, then all measurement series associated with the term *ship* should be returned as well as those associated with the more specific terms.

Although TRiDaS makes provision for controlled vocabularies, it does not include the vocabularies themselves. The construction of these vocabularies is a considerable undertaking that requires cooperation both within the dendrochronology and the wider academic community. Progress is being made in this regard by a number of groups, most notably the ARIADNE EU infrastructure project. A key goal of ARIADNE is to bring together and integrate archaeological infrastructures, including dendroarchaeological data through the Digital Collaboratory for Cultural Dendrochronolog (DCCD) repository. Once these vocabularies are established, they will be implemented within both DCCD and Tellervo.

The comprehensive nature of the TRiDaS format means that it is considerably more extensive than the formats that it aims to succeed. Although TRiDaS files remain human-readable and in theory could be written manually, in practice they need to be generated by applications and toolkits. The wide availability of support for TRiDaS in popular dendrochronological programs will be essential for the continued growth and adoption of TRiDaS. As of the time of writing, the applications capable of reading and/or writing TRiDaS files include Tellervo, TRiCYCLE (Brewer *et al.* 2011), TRiDaBASE (Jansma *et al.* 2012a), DCCD (Jansma *et al.* 2012b), and dplR (Bunn 2008). This article will focus on the use of the Tellervo application.

TELLERVO

Tellervo is a cross-platform desktop application for the measurement and curation of dendrochronological samples. It is open source and freely available for Mac, Windows, and Linux computers (http://www.tellervo.org). Tellervo also supports a wide variety of measurement platform hardware.

Tellervo has evolved from the original Cornell Ring Analysis System (Corina; Brewer *et al.* 2010) code base to be an enterprise-level server/client system primarily designed for medium to large dendrochronological laboratories. The server's secure web-service architecture means users both within the same laboratory and in different laboratories can access and work with data from the same database over the Internet. Tellervo uses TRiDaS at its core to provide users with the ability to describe and store a wide variety of data and metadata within a secure relational database.

TELLERVO CASE STUDY—YENIKAPI

The development of the Tellervo software has been carried out in parallel with the dendrochronological research arising from the extensive archaeological excavations at Yenikapı in Istanbul (Pearson *et al.* 2012). The needs of the dendroarchaeologists working at Yenikapı have therefore been instrumental in guiding the development of Tellervo.

The excavations at Yenikapı arose through the development of a major subway running under the Bosphorus to increase transport capacity in Istanbul. The excavation totals an area of 26,250 m^2 (Pekin and Kangal 2007) and covers what was the Byzantine port of Theodosius. The logistical challenge such an immense excavation poses not only to the archaeologists but to the dendrochronologists cannot be underestimated. At the time of writing, 2710 dendrochronological samples have been processed with several hundred more outstanding. Each of these has a wealth of information attributed to it in the field including location within the excavator's grid, structure from which it was taken, depth, dimensions, and various sample codes assigned by the excavators. It is essential that this information remains accessible to the dendrochronologists working on the wood.

Once the sample is analyzed in the lab, more information is generated that is essential for the interpretation of the samples. In addition to the raw ring-width measurements, this information includes species, presence/absence of pith, presence/absence of bark, presence/absence of sapwood, sapwood ring count, and unmeasurable rings. Tellervo provides the ability to record this information accurately and transparently within its relational database. Tellervo also offers the ability to record observations on a ring-by-ring basis. This feature has been useful for recording interesting features in the samples such as frost rings and other scars. Ring remarks can also be useful for recording practical information—for instance, which rings have been sampled for radiocarbon wiggle-matching analysis, which have pinholes (a standard method for marking decades on the physical sample), and the point at which a radius shift has been made (i.e. a point in a sample where the anatomy has forced the analyst to continue measuring elsewhere around the ring). Figure 1 illustrates the data and metadata entry tabs in Tellervo for one measurement series from Yenikapı. In this example, the data screen includes pinhole and radius-shift ring remarks illustrated by icons and the metadata screen is displaying some of the fields associated with the series.

The Yenikapı project has also benefited greatly from the introduction of a barcoding sample curation system within Tellervo. Following the initial data entry of samples into the system, Tellervo generates unique barcodes for each sample. These barcodes are fixed to the samples during the preparation stage and serve as a quick and effective way of identifying each sample within the system. When a user comes to measure a sample, a simple scan of the barcode retrieves all the relevant metadata from the database, which both speeds up the measuring process and minimizes user error through the mistyping of sample codes. The barcoding system has also proved invaluable in the curation of the samples. Samples are assigned to barcoded archival boxes, greatly simplifying the task of tracking the location of samples between the lab and archive rooms.

MAPPING AND SPATIAL ANALYSIS

The Tellervo database is spatially enabled through the use of PostGIS extensions to PostGreSQL within the Tellervo server package. The Tellervo desktop application makes good use of this by providing an integrated 3D mapping system capable of displaying site and sample locations at global through to local scales. All location information is tied directly to the database, allowing users to browse the map and link directly to the metadata and load ring-width series (see Figure 2 for a screenshot example). The map viewer can draw information from various online sources providing satellite and aerial imagery alongside standard maps. It also supports the reading of data from standard Web Map Services (WMS), providing a convenient option for laboratories that have additional spatial data that they would like to visualize alongside their dendro data. The map viewer also supports the display of local spatial data files in both ESRI Shapefile and Google Earth KML formats.

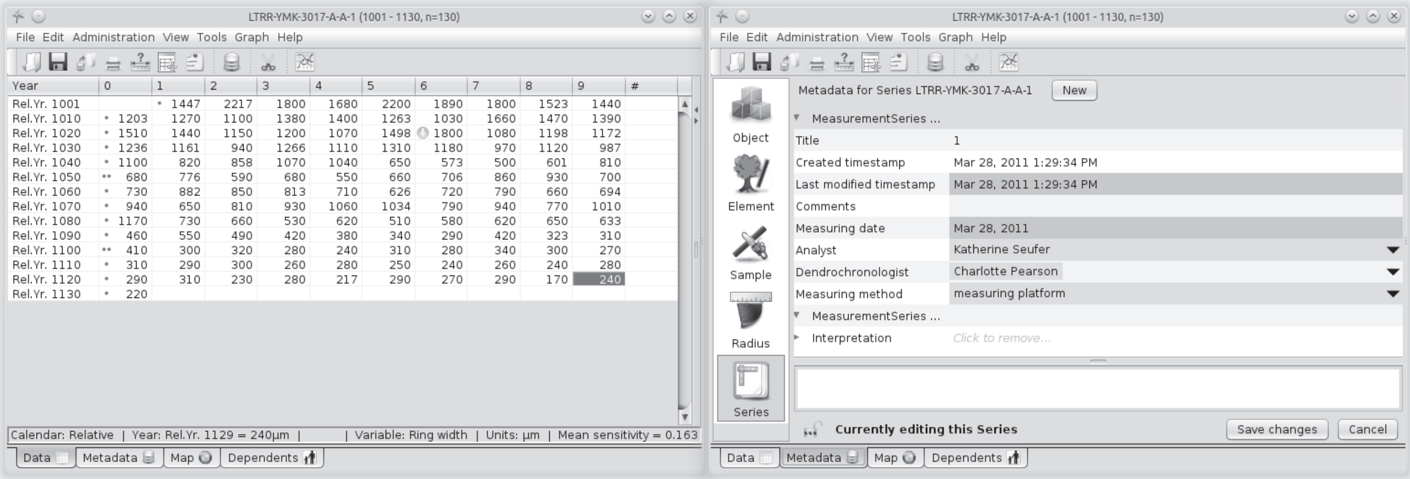

Figure 1. Screenshots of the data (left) and metadata (right) tabs within Tellervo for a measurement series from Yenikapı. The data tab shows the ring widths for a series that is not yet dated and includes a number of rings with ring remarks indicating the location of pinholes in the sample and the point at which a radius shift was necessary during measurement. The metadata tab shows some of the fields associated with this series. The object, element, sample radius, and series buttons on the left of the screen correspond to the TRiDaS entities. Selecting these entities shows the metadata fields associated with each.

Figure 2. Screenshot of the 3D mapping interface in Tellervo. The map shows a view south of the archaeological sites in Istanbul within this Tellervo database, overlaid on LandSat imagery of the city. Note the exaggerated topography, the map markers for sites sampled in and around the city, as well as the highlighted location of the Yenikapı excavation. The popup tag gives the user direct access to the metadata describing the site, as well as the ability to quickly access all associated data series.

The 3D map viewer capabilities within Tellervo are not intended as a fully fledged geographical information system. However, for users who would like to perform spatial analysis of their data, this can be done by connecting to the PostGIS-enabled Tellervo database from standard GIS software. One example of how this was used during the Yenikapı project is included in the publication by Pearson *et al.* (2012). They used the spatial data within Tellervo to explore phases of use and repair, in this case by plotting the location of samples within a specific dock. Although four phases were identified, the spatial distribution clearly showed that these were phases of repair rather than indicating a change of function associated with widening or lengthening of the dock.

CONCLUSIONS

The Tellervo software builds upon the success of the Tree-Ring Data Standard to provide users with a powerful yet user-friendly tool for dendrochronological research. Although Tellervo has been developed to address the needs of dendroarchaeologists, the inclusive nature of TRiDaS means it is suited for use in all spheres of dendrochronology.

Tellervo continues to be actively developed. As an open-source product, users are urged to contribute by providing feedback and suggestions for future development. If there are features you would like to see included in Tellervo, which will be useful to the wider community, please get in touch with the developers.

ACKNOWLEDGMENTS

The initial development of the Tellervo software was supported by the College of Arts and Sciences, Cornell University, and more recently by the Malcolm H. Wiener Foundation and the University of Arizona. The Yenikapı project would not have been possible without the long-term support of the National Science Foundation, the National Endowment for the Humanities, the National Geographic Society, the Malcolm H. Wiener Foundation, and numerous individual patrons of the Aegean Dendrochronology Project. Finally, I would like to thank two reviewers for their comments and suggestions for improvements to an earlier draft of this article as well as to the direction of the future development of Tellervo.

REFERENCES

Brewer, P., K. Sturgeon, L. Madar, and S. Manning, 2010. A new approach to dendrochronological data management. *Dendrochronologia* 28(2):131–134.

Brewer, P., D. Murphy, and E. Jansma, 2011. TRiCYCLE: A universal conversion tool for digital tree-ring data. *Tree-Ring Research* 67(2):135–144.

Bunn, A. G., 2008. A dendrochronology program library in R (dplR). *Dendrochronologia* 26(2):115–124.

Cook, E. R., B. M. Buckley, J. G. Palmer, P. Fenwick, M. J. Peterson, G. Boswijk, and A. Fowler, 2006. Millennia-long tree-ring records from Tasmania and New Zealand: A basis for modelling climate variability and forcing, past, present and future. *Journal of Quaternary Science* 21(7):689–699.

Friedrich, M., S. Remmele, B. Kromer, J. Hofmann, M. Spurk, K. F. Kaiser, C. Orcel, and M. Kuppers, 2004. The 12,460-year Hohenheim oak and pine tree-ring chronology from central Europe—A unique annual record for radiocarbon calibration and paleoenvironment reconstructions. *Radiocarbon* 46(3):1111–1122.

Fritts, H. C., 1963. Computer programs for tree-ring research. *Tree-Ring Bulletin* 25(3–4):2–7.

Grissino-Mayer, H. D., and H. C. Fritts, 1997. The International Tree-Ring Data Bank: An enhanced global database serving the global scientific community. *The Holocene* 7(2):235–238.

Grudd, H., K. R. Briffa, W. Karlén, T. S. Bartholin, P. D. Jones, and B. Kromer, 2002. A 7400-year tree-ring chronology in northern Swedish lapland: Natural climatic variability expressed on annual to millennial timescales. *The Holocene* 12(6):657–665.

Jansma, E., P. Brewer, and I. Zandhuis, 2010. TRiDaS 1.1: The tree-ring data standard. *Dendrochronologia* 28(2):99–130.

Jansma, E., R. van Lanen, K. Sturgeon, S. Mohlke, and P. W. Brewer, 2012a. TRiDaBASE: A stand-alone database for storage, analysis and exchange of dendrochronological metadata. *Dendrochronologia* 30(3):209–211.

Jansma, E., R. J. van Lanen, P. W. Brewer, and R. Kramer, 2012b. The DCCD: A digital data infrastructure for tree-ring research. *Dendrochronologia* 30(4):249–251.

Pearson, C. L., C. B. Griggs, P. Kuniholm, P. W. Brewer, T. Ważny, and L. Canady, 2012. Dendroarchaeology of the mid-first millennium AD in Constantinople. *Journal of Archaeological Science* 39(11):3402–3414.

Pekin, A. K., and Kangal, S., 2007. *Istanbul: 8000 Years Brought to Daylight (Marmaray, Metro, Sultanahmet Excavations)*. Vehbi Koç Foundation, Istanbul.

Salzer, M. W., and M. K. Hughes, 2007. Bristlecone pine tree rings and volcanic eruptions over the last 5000 yr. *Quaternary Research* 67(1):57–68.