Available online at www.sciencedirect.com
ScienceDirect
Dendrochronologia 28 (2010) 251–258
DENDROCHRONOLOGIA
www.elsevier.de/dendro

TECHNICAL NOTE

Statistical and visual crossdating in R using the dplR library

Andrew G. Bunn*

Environmental Sciences, Huxley College, Western Washington University, Bellingham, WA 98225-9181, United States

Received 26 September 2009; accepted 17 December 2009

Abstract

I demonstrate new functionality for the Dendrochronology Program Library in R (dplR) that allows for flexible statistical crossdating of tree-ring data. Using a well-dated ring-width file, I give examples of how dplR can be used to examine correlations between each series and a master chronology according to overlapping time periods (segments) specified by the user; examine moving correlations of suspect series; and compute cross-correlation functions to identify specific dating issues. I also show how automatically generated skeleton plots can be used to visually crossdate. Much of the terminology and approach used for crossdating in dplR will be familiar to users of COFECHA.

Keywords: Statistical software; Moving correlation; Cross-correlation; Skeleton plot; COFECHA

Introduction

The advent of computer-based tools has greatly advanced the field of dendrochronology. The Dendrochronology Program Library (Holmes, 1992) and ARSTAN (Cook and Holmes, 1996) brought new rigor to the field and are widely used by tree-ring scientists worldwide. The program COFECHA (Holmes, 1983) brought statistical crossdating into common practice and is employed by the International Tree-Ring Data Bank as the standard quality-control tool.

The Dendrochronology Program Library in R (dplR) (Bunn, 2008) is an open-source package used within the R statistical programming environment (R Development Core Team, 2009). The use of dplR makes it easier for dendrochronologists to use R as their primary analytic environment especially in conjunction with other tree-ring specific packages (e.g., bootRes; Zang, 2009). Here I describe functions that allow graphical crossdating using dplR version 1.3 (Bunn, 2010).

New crossdating functions and examples

I use Schulman's Mesa Verde Douglas Fir (*Pseudotsuga menziesii*) data from the International tree-ring data bank (Schulman, 1963), to demonstrate how dplR can be used to crossdate tree-ring series. The Mesa Verde data are exquisitely well dated with 35 series spanning the 13th century to the mid-20th century. The average series length is 565 years ($\sigma = 157$), autocorrelation at 1-year averages 0.60 ($\sigma = 0.16$), and mean sensitivity, calculated according to Eq. (2) in Biondi and Qeadan (2008), is 0.61($\delta = 0.11$). Each series is correctly dated and the average series correlation to the master chronology is 0.85 ($\sigma = 0.04$). These numbers, and indeed all the calculations and figures presented here, are reproducible using the code in Appendix A.

A common issue in crossdating is the presence of a missing ring. To demonstrate crossdating in R, I randomly chose a series from the Mesa Verde ring-width dataset and corrupted

*Tel.: +1 360 650 4252; fax: +1 360 650 7284.
E-mail address: andrew.bunn@wwu.edu.

1125-7865/$ – see front matter © 2010 Istituto Italiano di Dendrocronologia. Published by Elsevier GmbH. All rights reserved.
doi:10.1016/j.dendro.2009.12.001

it by deleting an existing measurement. The code below reads in the Mesa Verde data (included with dplR) and corrupts series 641143 by deleting the 325th measurement:

```
R>data(co021)

R>dat=co021

R>tmp=dat$'641143'

R>tmp=c(NA,tmp[-325])

R>dat$'641143'=tmp
```

The R object `dat` now contains the one misdated series. I will demonstrate how to locate the approximate location of that dating problem. First, the dplR function `corr.rwl.seg()` looks at the correlation of each series in a tree-ring dataset according to user-specified segment lengths and creates a plot where suspect series are flagged if the correlation to the master chronology is below a user-defined critical value. The correlation measure is Spearman's rank correlation coefficient (ρ) by default. In this function each series is removed from the dataset and a master chronology is calculated as the mean of the remaining series (using Tukey's biweight robust mean by default). Each series is prewhitened by default to remove autocorrelation using an autoregressive model where the parameters and complexity are selected by AIC. Users can opt to detrend each series using a Hanning filter to remove low-frequency variability prior to, or

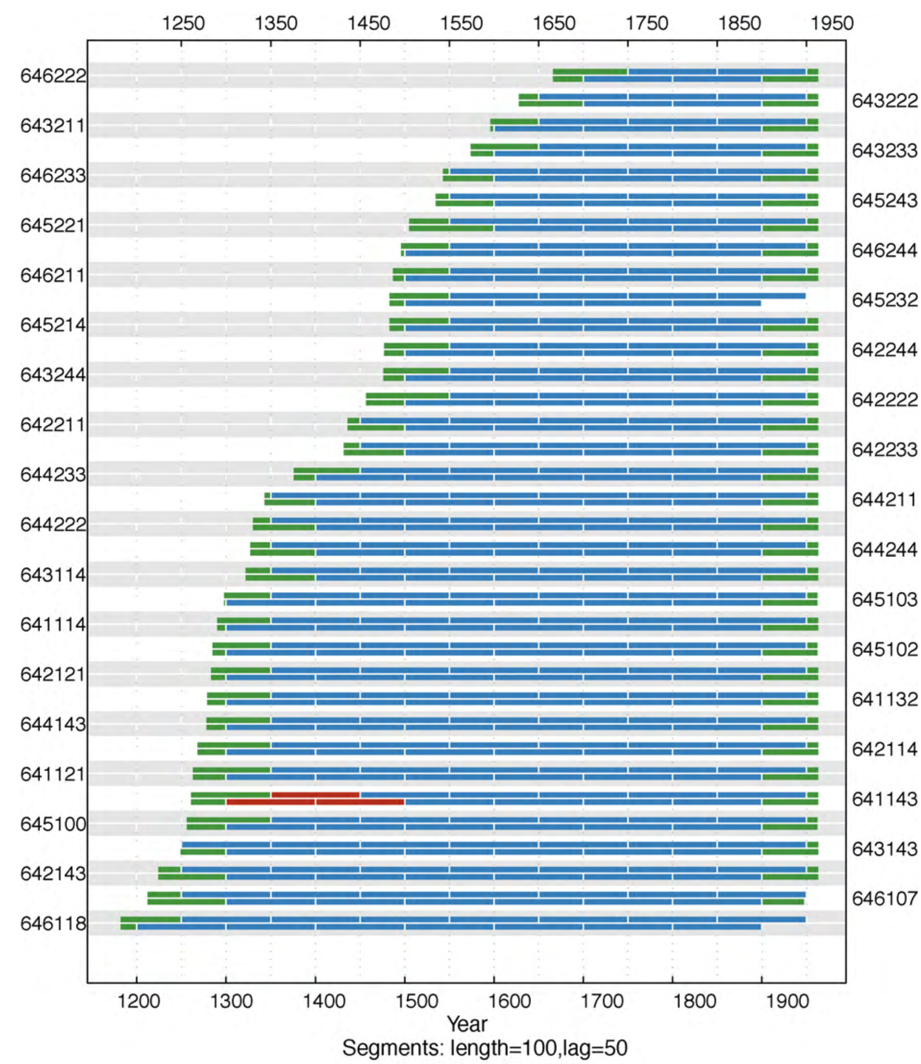

Fig. 1. Each segment of each series in Schulman's Mesa Verde dataset is shown and colored by correlation with the master chronology. Each series is represented by two courses of lines with the bottom course adhering to the bottom axis timeline and the top course matching the upper axis timeline (100-year segments lagged by 50 years). Segments are colored according to their correlation. Blue segments correlate well to the master chronology (*p*-values less or equal to the user-set critical value) while potential dating problems are indicated by the red segments (*p*-values greater than the user-set critical value). Green lines show segments that do not completely overlap the time period and for which no correlations are calculated. Series 641143 shows poor correlation beginning in the 16th century. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)