TECHNICAL NOTE

Exploration of long-term growth changes using the tree-ring detrending program "Spotty"

Jan Esper[a,b,*], Paul J. Krusic[c], Ken Peters[c], David Frank[a]

[a]*Dendro Sciences Unit, Swiss Federal Research Institute WSL, Zurcherstrasse 111, 8903 Birmensdorf, Switzerland*
[b]*Oeschger Centre for Climate Change Research, University of Bern, 3012 Bern, Switzerland*
[c]*Tree-Ring Laboratory, Lamont-Doherty Earth Observatory, Palisades, NY 10964, USA*

Abstract

A new tree-ring detrending program (Spotty) performs analyses of changing growth trends and environmental signals with tree age. Spotty is particularly useful in understanding the nature of long-term, centennial scale trends in tree-ring data but has a variety of dendrochronological applications. The program permits the user to define up to fifty age classes from a file of increments data. Age classes can be defined as ranges (e.g., 25–50, 100–150 years) or as single years (e.g., 25, 40, 55 years). The program will select the data corresponding to these classes and fit spline functions to the respective age class, data clouds. The user is allowed to choose a certain spline bandwidth, and can decide to fit splines to either the raw or Regional Curve Standardization (RCS) detrended data – with the RCS routine being included in Spotty – thus permitting the analysis of different underlying frequencies for each selected age class. The individual, age class splines may also be compared for common (or deviating) variance. Spotty facilitates color graphics and allows saving results for further use in other programs. PC and Mac versions of the program are available at the WSL Dendro Sciences (www.wsl.ch/forschung/forschungsunits/dendro) and LDEO Tree-Ring Lab (www.ldeo.columbia.edu/res/fac/trl/public/publicSoftware.html) web pages.
© 2008 Elsevier GmbH. All rights reserved.

Keywords: Tree rings; Age trend; Standardization; Detrending; Dendrochronology

Introduction

Tree-ring detrending methods are necessary to remove age-related trends in the radial growth rate (Cook and Kairiukstis, 1990; Fritts, 1976). Of the various approaches applied in dendrochronology, certain methods have proven to be particularly useful in retaining centennial scale variability. These include the Regional Curve Standardization (RCS; Mitchell, 1967; Briffa et al., 1992; Cook et al., 1995; Esper et al., 2003) and Age Banding (Briffa et al., 2001) methods, both resting on the basic concept of comparing tree rings of the same biological (cambial) age.

RCS involves the removal of the age trend archetype estimated from a collection of tree-ring data series and constructing a mean chronology of the entire data set. Age Banding selects data from a certain age class range (bands) and develops a chronology that is the average of the desired age classes. Age Banding thus allows the comparison of timeseries representing specific age class ranges, but the method can be restricted by low sample replication over time, challenging the development of a continuous timeseries.

---

*Corresponding author at: Dendro Sciences Unit, Swiss Federal Research Institute WSL, Zurcherstrasse 111, 8903 Birmensdorf, Switzerland. Tel.: +41 44 739 2510; fax: +41 44 739 2515.

E-mail address: esper@wsl.ch (J. Esper).

We developed a program "Spotty" that represents a continuation of these methods, as it relies on the same principle of associating equal-aged tree rings (Mitchell, 1967), but allows fitting spline functions to age classes ranging from single years to a single band representing the entire data set. The splines used in the program facilitate bridging data gaps and thus allow the estimation of a continuous timeseries from otherwise discontinuous data clouds. Time-varying error bars for all splines are computed and may be considered as an estimate of the uncertainty of the spline fit. Graphical comparison of the splines permits identification of varying long-term trends retained in the growth rates of different biological age classes.

In this note, we briefly describe the program Spotty and provide suggestions that may be useful in running the program. We demonstrate some features of the program using a data set that was previously analyzed to reconstruct large-scale temperature changes (Esper et al., 2002; herein abbreviated as ECS02).

Spotty design and application

Spotty is a top-down program where all parameters are defined first and then executed. This design permits continuous looping within the program allowing the user to change settings as often as desired. All program prompts are accompanied by brief descriptions or examples of acceptable input. In nearly every loop of the main program, intermediate results may be saved for post-execution display or analysis.

Currently Spotty is dimensioned to handle 2200 series, with a maximum length of 2100 years. The program can build up to 50 age classes at a time as long as no more than 4,620,000 values are provided. Single age classes or age class ranges defined by the user may contain all measurements for a single biological age, or all the measurements in a biological age range. The program allows the user to select age classes, extract these data from a file, plot those values over their calendar year of growth, and then fit splines to the resulting data clouds for interpretation.

After displaying any plot, Spotty provides an option for saving the plot as postscript or plot data as ASCII files. In addition, plot attributes (e.g., symbol color, axis limits) may be changed to emphasize desired elements in displayed figures. During program execution there is the option to begin again with a new chronology type, define new age classes, or exit the program.

Data types and organization

The essential input to Spotty is an increment file in decade format (http://www.ncdc.noaa.gov/paleo/treeinfo.html#formats). This format is able to be easily interpreted by humans and computers, and conveniently maximizes the number of measurements per line while minimizing the total number of lines per record. In addition Spotty will accept an ASCII file of pith offsets (i.e., estimates of the number of years missing to the pith) to more precisely align the tree-ring data by cambial age (Esper et al., 2003). The only requirements of this pith-offset file are that it must contain one title line at the top, and that the number of rows (pith-offset estimates) and their order be identical to the number of series in the increment (data) file. If no file of pith estimates is provided, the program assumes all measurements begin at the biological age of one.

For illustration, we use samples from living trees and relict wood integrating several species and sites from the Northern Hemisphere extratropics (Cook et al., 2004; Frank et al., 2007). These data were previously analyzed and split into a "linear" and "non-linear" data set based on the shape of tree-age-related growth trends (Esper et al., 2002). Chronologies from these data sets were developed using RCS, and the mean timeseries back-transformed into mm units to highlight stem growth variations, which can potentially be associated with tree biomass (Fig. 1). As the average age/growth of the non-linear data is lower/higher (195 years/0.60 mm) than that of the linear data (300 years/0.42 mm), the rescaled RCS timeseries from the non-linear data is mostly above its linear counterpart. Besides this difference in stem productivity, the data indicate higher growth rates during Medieval times, reduced growth during an extended period ∼1300–1800, and increasing growth towards present.

Using Spotty, these long-term changes are here assessed by analyzing the constituent growth trends retained in the various age classes of the underlying data (Fig. 2). For this purpose, the user may choose to work with one of three different variables selected from the Data-Transformation-Menu. The choices are: (i) the original raw data (Fig. 2B), (ii) smoothed estimates of the original raw data (Fig. 2C) or (iii) RCS detrended data (Fig. 2D). Methodological variations of RCS detrending, such as the degree of smoothing for the regional curve construction and the choice of computing ratios or residuals (Esper et al., 2003; Büntgen et al., 2005), can be chosen within the program. While all subsequent analyses shown here are performed on the raw ECS02 measurements (Fig. 2B), the data input type can be changed by revisiting the Data-Transformation-Menu.

Age class data clouds

The next window is the Main-Program-Menu from which desired data clouds are built. A data cloud is the
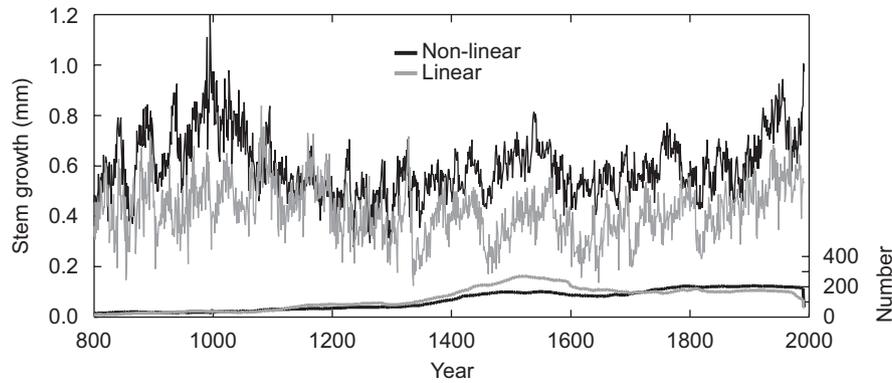
Fig. 1. RCS chronologies of trees classified as "linear" and "non-linear" based on the shape of their growth trends from ECS02. Chronologies are expressed as records of radial stem productivity (in mm) of high elevation and high latitude NH forests since AD 800. Curves at the bottom show the number of ring width measurements integrated in the chronologies.

graphical representation of all measurements of one or more user-defined biological ages or age ranges. For example, the data cloud of age class 1 consists of all measurements from trees containing values for cambial age 1. A data cloud of the age class 100–200 will contain only values from trees in their 100th–200th years of growth. The program can plot and analyze up to 50 of these data clouds at a time allowing one to compare tree growth for multiple age classes.

There are three different methods by which the user may build a single or multiple age class definitions: (i) Random, (ii) Automatic or (iii) Manual. The Random age class definition routine requires a start and stop age, the number of classes to build within the start and stop age-span, and a random number seed. Spotty will then make the desired number of contiguous age classes, of random length, between the start and stop ages. The Automatic mode requires a start and stop age for the youngest age class, then an increment by which to define the next age class, and finally the maximum age of the last class. In Manual mode, the program first prompts for the total number of age classes to build, and then the start and stop ages of each. Values used to select data shown in Fig. 3 were: ⟩1⟨ (start), ⟩1⟨ (stop), ⟩100 ⟨ (increment), ⟩550⟨ (maximum); automatic mode.

Splines and error

Once age classes are defined and the corresponding data selected, the program fits a user specified spline to each cloud to emphasize age class specific growth behavior. The variability of the splines is, however, not only bandwidth defined, but also changes with the number of values available in each age class. This is highlighted in the 100-year splines fit to the 1, 101, 201, 301, 401, and 501 clouds that appear less flexible as fewer data points are available (Fig. 3). In addition to

the splines commonly used in dendrochronology (Cook and Peters, 1981), Spotty offers the application of a generalized cross validation [gcv] spline (Hutchinson and De Hoog, 1985) that attempts to "automate" the bandwidth selection by minimizing the gcv statistic (see the Appendix).

Besides highlighting specific data clouds and spline fits, Spotty also provides an estimate of the error of the spline fit (see the Appendix). The error is generally smaller with stiffer splines, as seen for the 10- and 100-year splines fit to age class 1 of ECS02 (Fig. 4). Naturally the error is also larger during periods that contain little data, such as the beginning of the 11th century or the end of the 12th century in Fig. 4A. Estimates of spline uncertainty can be employed for spline selection, further refinement of age class selection, and as an aid in interpretation.

Common variance analysis

A combined illustration of data clouds and their splines – shown with changing colors and symbol types in Spotty – helps identify common (and uncommon) lower frequency signals in different aged tree rings. For the ECS02 data used here, a trend of decreasing spline variability towards the older (and less replicated) age classes, as well as a trend of decreasing ring widths with aging, i.e. splines are ordered by age (Fig. 5A), is visible. The latter can be removed by centering or normalizing the data clouds and their splines, by removing the common period (or some user-defined) mean, or both the common period mean and standard deviation as in Fig. 5B.

The normalized data clouds and splines in Fig. 5B indicate changing long-term growth trends with aging. The splines fit to the young age classes are characterized by generally higher growth values at around 1500, their
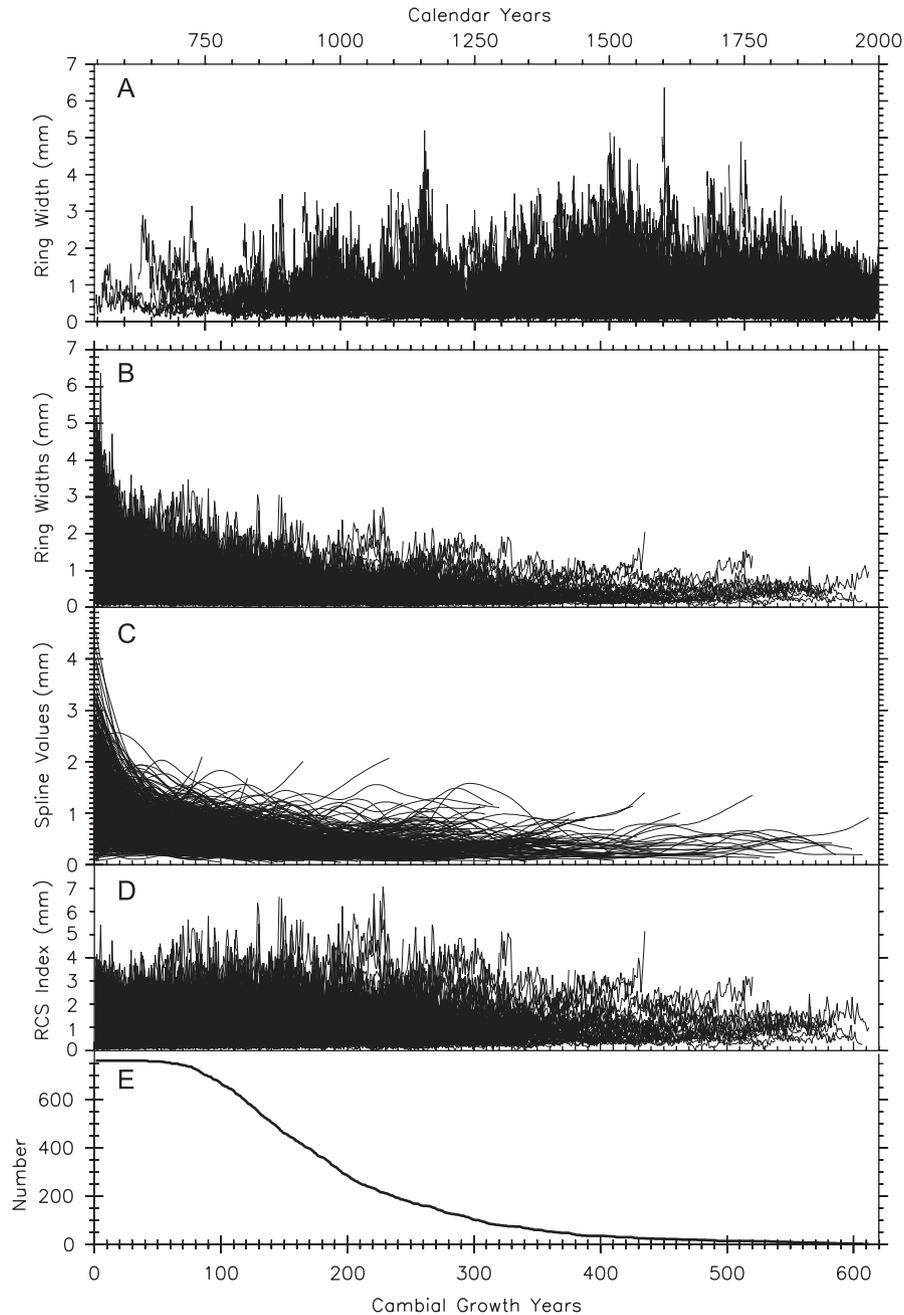
Fig. 2. Non-linear ECS02 data aligned by calendar years (A), and cambial growth years (B–D). (E) The sample replication of the age-aligned data. Spotty allows working with raw (B), smoothed (C), and RCS detrended (D) data. For RCS detrending, a 10-year spline was fit to the age-aligned data, and residuals from this regional curve calculated.

counterparts fit to older age classes show increased growth values at the beginning and/or towards the end of the past millennium, though old classes are represented by less data points and appear less flexible.

An alternative comparison of splines – this time fit to the RCS detrended ECS02 data (RCS detrending effects are similar to data cloud normalization) – supports the finding from the previous experiment (Fig. 6). Splines fitted to the "younger" age classes 1–151 and 161–241 tend to show more decadal scale variance – including

some opposing deviations in the 12th century – but less centennial scale trends. The "older" and less replicated age classes 251–361 and 371–491 are characterized by a different long-term behavior including synchronous centennial scale trends and high growth values at the beginning and end of the past millennium. The reduced long-term variability in the youngest age classes perhaps suggests that low frequency variance is less common in these data. This could be due to, for example, increased mechanical stresses adding noise to the ring width
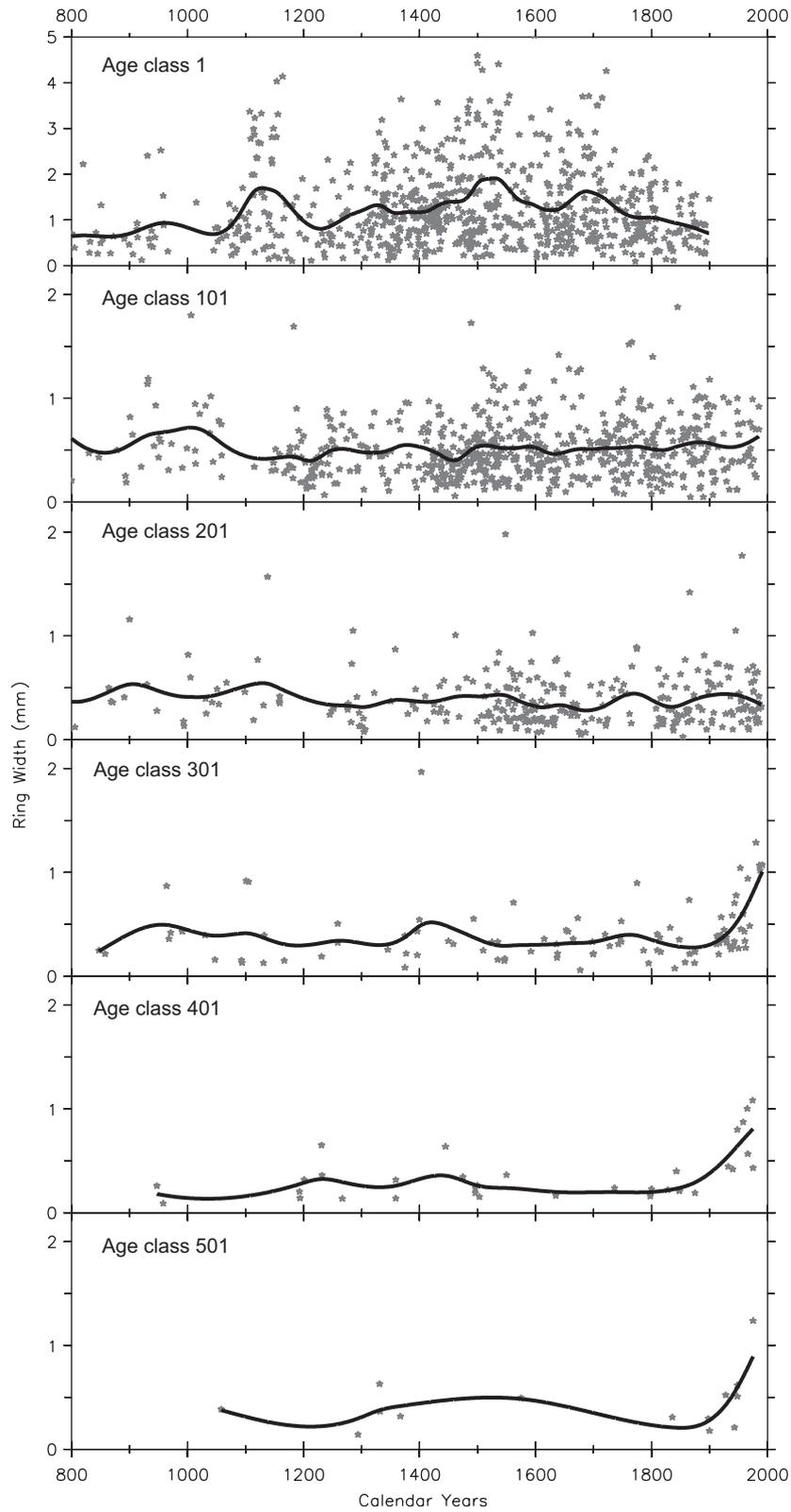
Fig. 3. Data clouds of age classes 1, 101, 201, 301, 401, and 501 of ECS02. Curves are 100-year splines.

data from young and small trees (Esper et al., 2008), and systematically poorer fits to the juvenile rings particularly when not utilizing pith-offset data. The older age class splines appear to be influenced by the lower sample replication, reduced flexibility, and increased error.
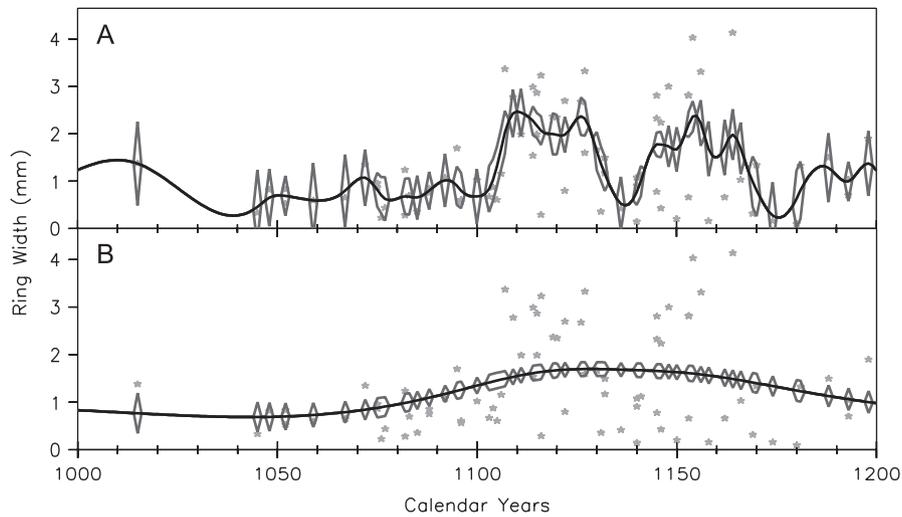
Fig. 4. Spline error estimates for age class 1 of the non-linear ECS02 data for (A) 10-year and (B) 100-year splines, both over the 1000–1200 period.
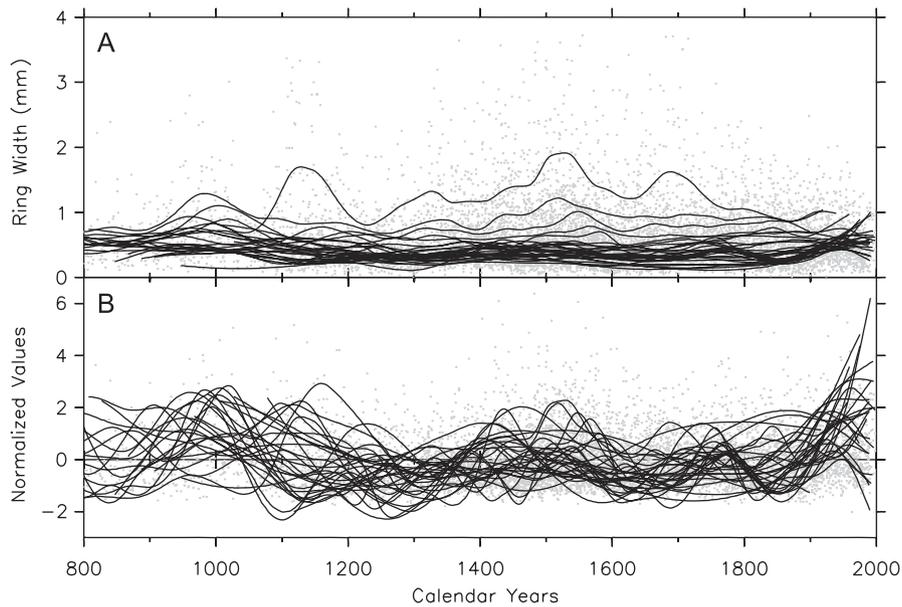
Fig. 5. (A) Value distributions of age classes 1, 21, 41, ..., 581 in mm, together with the 100-year splines fit to these data. (B) Same splines, but normalized by removing both the means and standard deviations over the maximum individual series lengths (max. 800–1990).

Calculations of 2-year splines in Spotty, this time fitted to larger age ranges 1–150, 151–240, and 241–500 years (Fig. 6C), allows a more detailed assessment of the changing variance stored in these data portions. These splines indicate, for example, that the increased growth values in the 12th century are restricted to the middle (151–240) age class. The positive deviations around 1400 and in the late 18th century are recorded in the middle and old (241–500) age classes. Also the recent increase since about the middle of the 19th century is stronger in the old tree-ring data >240 years.

Conclusions

Spotty enables analyses of common variance in tree-ring data as a function of biological age. It allows the user to define age classes from a file of increment data, then fit a spline to the cambially adjusted age class data clouds for analysis of underlying, age-related, low frequency variability. The program might be particularly useful to analyze lower frequency trends in tree-ring data, and can be used to identify data fractions that might add uncommon, perhaps biasing variance to
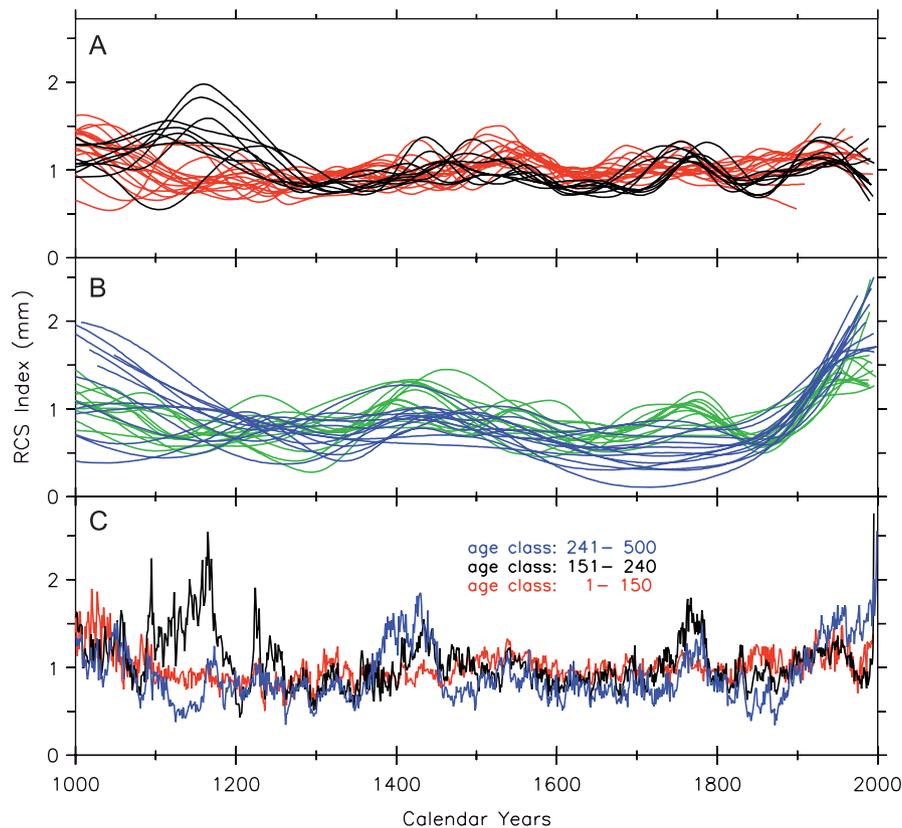
Fig. 6. Hundred- and 2-year splines fit to age classes (years and ranges) of the RCS detrended non-linear ECS02 data. (A) Hundred-year splines fit to age classes 1, 11, …, 151 in red, and to age classes 161, 171, …, 241 in black. (B) Hundred-year splines fit to age classes 251, 261, …, 361 in green, and to age classes 371, 381, …, 491 in blue. (C) Two-year splines fit to age class ranges 1–150 (red), 151–240 (black), and 241–500 (black). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mean chronologies (Esper et al., 2007). Spotty may be readily applied in ecological studies where detrending growth increments is not desirable and where one may wish to avoid the confounding effects of tree age (Rolland et al., 1998; Sarris et al., 2007). An improved understanding of tree-age-related variability and signal changes should be useful in a variety of dendrochronological related applications.

Acknowledgements

Appendix

URL: Spotty can be downloaded at

www.wsl.ch/forschung/forschungsunits/dendro/
www.ldeo.columbia.edu/res/fac/trl/public/
publicSoftware.html

Splines: In Spotty splines can be fit to individual tree-ring sequences to calculate indices or residuals, and to data clouds of specified age classes for comparative analyses of common variance. The user is allowed to specify a bandwidth for each data sequence or to have the program find a minimum generalized cross validation [abbreviated *gcv* in Spotty] spline.

The *gcv* spline has been discussed by Wahba (1983) and others as an objective way to select a stiffness parameter for fitting a smoothing spline to noisy data. *Gcv* is a statistic that is computed from the fit. It varies smoothly with the bandwidth and should have a well-defined and unique minimum for a given data set. This minimum value and the corresponding bandwidth define the minimum *gcv* spline through the data and it is thought to be in some sense optimal (details in Hutchinson, 1986; Wahba, 1983).

Fit statistics: In cases where a spline is fit to a selected age class, the data may need to be modified so that there is only one data point (raw or otherwise) for each year for which there is data, and for each such year a weight is computed as $1/sqrt(n)$, where n denotes the number of data points for that year. Thus the number of points splined [abbreviated *nps* in Spotty] for a cloud will be

not more than the total number of data points in a cloud (usually less), and the mean squared weight of the splined data points [*msw*] will be not more than one (also usually less). These weights ensure a closer fit for averages based on more than one cloud point. For a single tree-ring sequence these statistics are both one. The mean square residual [*msr*] to the fit is computed as usual and its square root gives a measure of average spread of the data points about the spline curve. The degrees of freedom [*df*] associated with the mean square residual and other statistics depends on the bandwidth. It is zero for a spline that fits the splined points exactly (an interpolating spline), equal to *nps* for a straight line fit, and in general increases smoothly with bandwidth. The square root of the mean square error [*MSE*] of the fit is an estimate of the average error associated with the spline curve itself. If it is added to the *msr* the result [*EV*] is an estimate of the average total squared error of using the spline curve as the "true" ring width for a particular year. EV is used to estimate the standard errors of the spline curve at each year.

Error: The errors depend on the bandwidth in at least two ways. The smoothing parameter, which is computed from the bandwidth, is a constant scale factor for all the errors. It is smaller for larger bandwidth. Also, the particular value of the bandwidth determines the localness of the spline smoothing. A small bandwidth results in a flexible spline and larger and more variable errors because fewer data points are used to compute both the fit and the errors. The errors also depend on the data density, being larger near gaps in the data, and in particular near the ends of the data. Finally, the errors also depend on the weight, which is assigned to the composite data point for each year, which is $1/sqrt(n)$. A bandwidth of 100 years for 200 years of data is not extreme, so the errors should vary some, but smoothly.

References

Briffa, K.R., Jones, P.D., Bartholin, T.S., Eckstein, D., Schweingruber, F.H., Karlén, W., Zetterberg, P., Eronen, M., 1992. Fennoscandian summers from AD 500: temperature changes on short and long timescales. Climate Dynamics 7, 111–119.

Briffa, K.R., Osborn, T.J., Schweingruber, F.H., Harris, I.C., Jones, P.D., Shiyatov, S.G., Vaganov, E.A., 2001. Low-frequency temperature variations from a northern tree ring density network. Journal of Geophysical Research 106, 2929–2941.

Büntgen, U., Esper, J, Frank, D.C., Treydte, K., Schmidhalter, M., Nicolussi, K., Seifert, M., 2005. The effect of power transformation on RCS – evidence from three millennial-length alpine chronologies. In: Gärtner, et al. (Eds.), Proceedings volume. International Conference on Tree Rings in Archaeology, Climatology and Ecology, April 22–24, 2004, Birmensdorf, Switzerland. Environment 53, Jülich, pp. 141–149.

Cook, E., Kairiukstis, A. (Eds.), 1990. Methods of Dendrochronology: Applications in the Environmental Sciences. Kluwer, Dordrecht, 394pp.

Cook, E.R., Peters, K., 1981. The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. Tree-Ring Bulletin 41, 45–53.

Cook, E.R., Briffa, K.R., Meko, D.M., Graybill, D.S., Funkhouser, G., 1995. The 'segment length curse' in long tree-ring chronology development for paleoclimatic studies. The Holocene 5, 229–237.

Cook, E.R., Esper, J., D'Arrigo, R., 2004. Extratropical Northern Hemisphere temperature variability over the past 1000 years. Quaternary Science Reviews 23, 2063–2074.

Esper, J., Cook, E.R., Schweingruber, F.H., 2002. Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. Science 295, 2250–2253.

Esper, J., Cook, E.R., Krusic, P.J., Peters, K., Schweingruber, F.H., 2003. Tests of the RCS method for preserving low-frequency variability in long tree-ring chronologies. Tree-Ring Research 59, 81–98.

Esper, J., Frank, D.C., Büntgen, U., Verstege, A., Luterbacher, J., Xoplaki, E., 2007. Long-term drought severity variations in Morocco. Geophysical Research Letters 34.

Esper, J., Niederer, R., Bebi, P., Frank, D.C., 2008. Climate signal age effects-evidence from young and old trees in the Swiss Engadin. Forest Ecology and Management 255, 3783–3789.

Frank, D., Esper, J., Cook, E.R., 2007. Adjustment for proxy number and coherence in a large-scale temperature reconstruction. Geophysical Research Letters 34.

Fritts, H.C., 1976. Tree Rings and Climate. Academic Press, New York, 567pp.

Hutchinson, M.F., 1986. A fast procedure for calculating minimum cross validation cubic smoothing splines. ACM Transactions on Mathematical Software 12, 150–153.

Hutchinson, M.F., De Hoog, F.R., 1985. Smoothing noisy data with spline function. Numerische Mathematik 47, 99–106.

Mitchell, V.L., 1967. An investigation of certain aspects of tree growth rates in relation to climate in the central Canadian boreal forest. Department of Meteorology, University of Wisconsin, Technical Report, 33pp.

Rolland, C., Petitcolas, V., Michalet, R., 1998. Changes in radial tree growth for *Picea abies*, *Larix decidua*, *Pinus cembra* and *Pinus uncinata* near the alpine timberline since 1750. Trees 13, 40–53.

Sarris, D., Christodoulakis, D., Körner, C., 2007. Recent decline in precipitation and tree growth in the eastern Mediterranean. Global Change Biology 13, 1187–1200.

Wahba, G., 1983. Bayesian "confidence intervals" for the cross-validated smoothing spline. Journal of the Royal Statistical Society: Series B 45, 133–150.