Appendix B. Tree-ring Standardization and Flow Reconstruction

This appendix contains details on standardizing of the tree-ring chronologies and converting them into statistical estimated of annual flow. An initial section titled "Standardization" is followed a section titled "Reconstruction Method." Literature cited is listed at the end of the appendix.

Standardization

Standardization refers to the procedure by which time series of measured ring widths of individual radii of trees at a site (usually >15 trees) are combined into a site chronology (Fritts, 1976). The site chronology is a single time series, dimensionless, with a mean of 1, such that an index greater than 1 means greater than normal growth and an index less than 1 means lower than normal growth.

Standardization was accomplished in this study with Matlab functions written by David Meko. The standardization approach is broadly similar to that taken in other streamflow reconstructions (e.g., Meko et al, 2001, 2007), and is similar to that used in the ARSTAN computer package (Cook et al., 2007). The main dedicated Matlab functions for standardization here are treeprep_az2 and sitechron1 The procedure is summarized in the numbered list below.

1. Organize the input files of measured ring width from the 69 sites. The following steps are then repeated for each site.
2. Fit each ring width series at the site with a cubic smoothing spline (Cook and Peters, 1981) that has an amplitude of frequency response of 0.95 (almost perfect tracking) at a wavelength twice the length of the measured ring-width series. This is very inflexible spline which is conservative in design in that it removes only very gradual trend, which cannot be distinguished from changes in ring width due to age or size of tree. As a consequence, very low frequency (e.g., multi-century) climate variations, should they exist, would not be retained in the detrended series.
3. Compute the standard core index as the ratio of the measured ring width to the value of the fitted spline. This is called ratio-indexing, and converts measured widths to a dimensionless core tree-ring index with a mean of 1.0 representing "normal" growth.
4. A precaution must be taken in ratio indexing to guard against an "exploding" index (Cook et al., 2007), which can result if the fitted spline curve approaches zero (division by 0 is infinity). Our approach is as follows. If the fitted spline at any drops below equal to or below the minimum measured ring width $w_{\min}$, we use only that part of the core index time series with the longest consecutive sequence of years of index greater than $w_{\min}$. It turns out that this truncation comes into play for only 8 of the 69 chronologies. And for those, no more than two cores are affected.
5. Compute residual versions of the core indices using autoregressive modeling. The objective is to remove persistence, which can vary greatly in standard indices from tree to tree and is often much greater than the persistence in the climate or hydrologic series being reconstructed. An order 1-3 autoregressive (AR) model is fit to each standard core index time series, and the autoregressive residual (with mean restored to 0) is defined s the residual core index (Cook and Kairiukstis, 1990). The order of the AR model is selected using an adjusted Akaiki Information Criterion, or AIC (Hurvich and Tsai (1989).
6. Average the standard core indices from all available cores together to compute the standard site chronology, and likewise for the residual site chronology. Following recommendation of Osborne et al. (1997), core indices are scaled to equal variance before averaging.

7. Summarize the common signal in the standard and residual site chronologies in terms of the mean between-series correlation, $\bar{r}$, of the core indices (Cook and Kairiukstis, 1990) and the expressed population signal (EPS; Wigley et al., 1984), and record the year the EPS first exceeds 0.85. EPS is a function of the sample size (number of cores) and $\bar{r}$, such that high $\bar{r}$ and high sample size favor a high EPS. As a rule of thumb, when EPS>0.85 the sample is large enough to represent the unknown population tree-ring signal at the site. We use this information to decide when to truncate the site chronology on the early end in reconstructing or inferring climate.
8. Stabilize the variance of the site chronology. The detrending operation described in steps 2-3 above refers to detrending of the mean (the average over core indices). The variance of the site chronology may also change gradually in unknown ways over time as the trees age and increase in size. The variance of the site chronology would also be expected to decrease systematically with increasing sample size because averaging tends to reduce noise, whose variance is not shared across trees. We adopted the method suggested by Osborn et al. (1997) to adjust for variance change over time due to change in sample depth. By this method, the departures of the site chronology are effectively scale according to a time-varying "independent sample size" that is a function of the number of cores and the mean between-core correlation of core indices.

Reconstruction Method

As described in the paper, the reconstruction method is modified from similar methods used in published reconstructions of precipitation and streamflow (Meko, 1997; Meko et al., 2001, 2007, 2011). Two reconstructions of annual flow are generated – a short reconstruction (1640-2001 CE) and a long reconstruction (903-2008 CE). Reconstruction steps are the same for both. Many tree-ring chronologies that contribute strongly to the short reconstruction do not extend back in time enough to cover the target period of the long reconstruction.

The reconstruction process for this paper can broadly be divided into two stages. First is single-site reconstruction (SSR) to convert each tree-ring chronology into an estimate of the target annual flows. Second is multi-site reconstruction (MSR) to combine the individual SSRs into a single final reconstruction. The SSR step needs to be done only once for each tree-ring chronology. The MSR step is done once for the short reconstruction and once for the long reconstruction. As described in the paper the regression models that provide reconstructed flows are actually calibrated on the square root of the water-year-total annual flows. These reconstructed flows are back-transformed to original flow units for plotting and for quantifying dry periods and wet periods.

All reconstruction steps were done within the Matlab programming environment using scripts and functions written by David Meko. The driving script that calls various reconstruction functons is Recon1_az2.m. Steps in reconstruction are listed below, roughly in their chronological sequence for the reconstruction process. For brevity, Q denotes annual flow in original units (BCM; billions of cubic meters, water year) and y is the square-root-transformed flow, or $y = \sqrt{Q}$. We used what is call the "residual chronology" (autoregressive residual; Cook and Krusic, 2007) in our reconstructions, and refer to this from now on as the "chronology."

1. Single site reconstruction (SSR) – repeated for each tree-ring chronology
    a. Stability check. If a difference of correlations test (Panofsky and Brier, 1968; Snedecor and Cochran, 1989) shows a significant $(p < 0.05)$ difference in correlation of y with the chronology in the first versus last halves of the overlap period, the SSR for the chronology is flagged to omitted from use in later steps.
    b. A pool of 10 potential predictors is set up: the chronology and its square at lags $t-2$ to $t+2$ years relative to y in year t.
    c. Preliminary stepwise regression (Weisberg, 1985) using cross-validation to choose how many steps are justified in the SSR regression model (selection of cutoff step). Leave-9-out cross-validation is used in this modeling to ensure independence of the calibration and validation data in the presence of lagged predictors (Meko, 1997).
        i. At each iteration of the leave-9-out regression modeling, store the value of the of the cross-validation prediction and the "deleted" residual (Weisberg, 1985). By this process time series of cross-validation predictions and deleted residuals are produce. Repeat p for steps $1, \ldots, m$, where m is the last step before the model stops entering or removing predictors according the *p*-enter and *p*-remove criteria (0.25,0.50). These criteria are intentionally lenient to allow the model to run an excessive number of steps.
        ii. At end of end of the iterations, use the m time series of deleted residuals to compute the reduction of error (RE; Fritts et al., 1990). Plot RE and the root-

mean-square error of validation residuals (RMSE_{cv}) as a function of step $j=1,\ldots,m$.

   iii. Choose as the stopping step for stepwise regression the step before the first drop in RE (or equivalently, the step before the first rise in RMSE_{cv}
   iv. For the special case of no significant predictors (no predictors enter) choose as the single predictor among the 10 potential predictors the one most highly correlated with y. Flag such a chronology to be skipped in later reconstruction steps. The chronology essentially has no signal for flow. Essentially, and SSR is generated for information purposes, but is not used later steps of reconstruction.

d. Recalibrate the model using all observations and specifying that the stepwise process stop at the step arrived at in (3) above
e. Store the calibration and leave-9-out cross-validation statistics for the selected model
   i. Calibration: calibration period, predictors in model; overall-F, $p-$value of F, R^2, R^2_{adj} (statistics defined in general regression texts (e.g. Weisberg, 1985))
   ii. Validation: RE, RMSE_{cv}, and r^*, the Pearson correlation between the reconstruction and the time series of cross-validation predicted values (see note at end of this appendix)
f. Plug the long-term tree ring data into the fitted regression model to get the SSR of y from each chronology
g. Repeat steps 1.a – 1.f for all chronologies
h. Flag the SSRs to be dropped. If any of the following five conditions are true, a chronology's SSR is not used in the subsequent step of multi-site reconstruction.
   i. The chronology fails the difference of correlation (first half vs second) test; in other words, has a temporally unstable signal (see step 1.a above)
   ii. Stepwise regression indicates that no variables from the pool of potential predictors enters the model (1.c.iv above)
   iii. $p-\mathrm{value}>0.05$ for overall$-F$ of regression (weak signal)
   iv. $\mathrm{RE}\leq 0$ from full cross-validation (no skill of model applied to independent data)
   v. p>0.05 for one-tailed test of significance* of correlation of observed flow with leave-9-out cross-validation predictions
i. Store SSRs in a time series matrix for subsequent use in multi-site reconstruction

2. Multi-site reconstruction (MSR); this sequences of steps run first for set of SSRs with time coverage for long reconstruction, then for set of SSRs with time coverage for short reconstruction.
   a. Average over the available SSRs (those not dropped) to get a single time series, w
   b. Regress annual flows y on w in simple linear regression using as a calibration period the full available overlap of the two series; the result is the reconstruction equation $\hat{y}=\hat{a}+\hat{b}w$, where w is the arithmetic mean of SSRs, $\{\hat{a},\hat{b}\}$ are the estimated regression coefficients, and $\hat{y}$ is the multi-site reconstruction, or MSR
   c. Run leave-9-out cross-validation of the estimated regression model
   d. Store calibration and validation statistics (as for SSRs described above)
   e. Substitute values of the SSRs before the start of the calibration period to get the full-length reconstructed time series of $\hat{y}$.

f. Convert reconstruction back to original flow units for plotting and evaluating severity of dry and wet years and intervals

****Significance of correlation of observed flows with cross-validation predictions***.

To avoid relying on assumptions of normality and non-autocorrelation in the test of significance of the correlation r between time series of observed flows y and leave-9-out cross-validation predictions $\hat{y}_{cv}$, we use Monte Carlo simulation. We generated many (10,000) simulations of $\hat{y}_{cv}$, by the method of exact simulation (Percival and Constantine, 2006), correlated each of the simulations with y, and compared the single correlation of the $\hat{y}_{cv}$ and y with the cumulative distribution function (CDF) of the correlations for the simulations.

Let p be the non-exceedance probability of that single r interpolated from the empirical CDF. For a one-tailed test for positive correlation, where ρ is the population correlation, H0 is $\rho \le 0$, H1 is $\rho > 0$, and the p-value for rejecting H0 was estimated as $1-p$. For a two-tailed test, H0 is $\rho = 0$, H1 is $\rho \ne 0$, and the estimated p-value is

$$p\text{-value} = \begin{cases} 2(1-p); & p \ge 0.5 \\ 2p; & p < 0.5 \end{cases}$$

The CDF for the above estimate is computed by these steps: 1) rank the N correlations of simulations and observed flows from smallest to largest $(i = 1, \ldots, N)$, and 2) assign each ranked correlation a probability $p_i^* = i/(N+1)$ of not being exceeded. The p-value, or the non-exceedance probability for the observed r_{xy}, is then interpolated linearly from a lookup table with columns $\{i \quad p_i^*\}$.

Literature Cited

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

Cook, E. R., P. J. Krusic, R. H. Holmes, and K. Peters (2007), Program ARSTAN,Version 41d, 2007, www.ldeo.columbia.edu/tree-ring-laboratory.

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

Fritts, H. C. (1976), Tree Rings and Climate, Academic Press, London, 567 pp.

Fritts, H. C., J. Guiot, and G. A. Gordon (1990), Verification, in Methods of Dendrochronology: Applications in the Environmental Sciences, edited by E. R. Cook and L. A. Kairiukstis, pp. 178–185, Kluwer Academic Publishers.

Hurvich, C. M., and C.-L. Tsai (1989), Regression and time series model selection in small samples, Biometrika, 76(2), 297–307, doi:10.1023/A:1005314315270.

Meko, D. (1997), Dendroclimatic reconstruction with time varying predictor subsets of tree indices, J. Climate, 10(4), 687–696.

Meko, D. M., D. W. Stahle, D. Griffin, and T. A. Knight (2011), Inferring precipitation-anomaly gradients from tree rings, Quatern. Int., 235, 89–100.

Meko, D. M., M. D. Therrell, C. H. Baisan, and M. K. Hughes (2001), Sacramento River flow reconstructed to A.D. 869 from tree rings, J. Am. Water Resour. Assoc., 37(4), 1029–1040.

Meko, D. M., C. A. Woodhouse, C. H. Baisan, T. Knight, J. J. Lukas, M. K. Hughes, and M. W. Salzer (2007), Medieval drought in the Upper Colorado River Basin, Geophys. Res. Lett., 34(L10705), 10.1029/2007GL029,988.

Osborn, T. J., K. R. Briffa, and P. D. Jones (1997), Adjusting variance for sample-size in tree-ring chronologies and other regional mean timeseries, Dendrochronologia, 15, 89–99.

Panofsky, H. A., and G. W. Brier (1968), Some Applications of Statistics to Meteorology, The Pennsylvania State University, University Park, Pennsylvania, 224 pp.

Percival, D. B., and W. L. B. Constantine (2006), Exact simulation of Gaussian time series from nonparametric spectral estimates with application to bootstrapping, Statistics and Computing, 16, 25–35.

Snedecor, G. W., and W. G. Cochran (1989), Statistical Methods, eighth ed., Iowa State University Press, Ames, Iowa, 503 pp.

Weisberg, S. (1985), Applied Linear Regression, 2nd ed., John Wiley, New York, 324 pp.

Wigley, T. M. L., K. R. Briffa, and P. D. Jones (1984), On the average value of correlated time series, with applications in dendroclimatology and hydrometeorology, J. Clim. Appl. Meteor., 23, 201–213