This file describes files output by TRISH for reconstruction by method MLR1-PCA (Multiple Linear Regression 1, on Principal Components). This is a two-stage method for reconstruction of a hydologic time series, *y*, from a network of tree ring chronologies, X, using stepwise multiple linear regression of *y* on principal components (PCs) of single-site reconstructions (SSRs) of *y*. A similar approach has previously been used in reconstruction of Colorado River annual flows (Meko et al. 2007). The two-stage approach to dendroclimatic reconstruction in general was introduced in the context of precipitation reconstruction by Meko (1997).

The first stage of the reconstruction method is to convert individual tree-ring chronologies to single-site reconstructions (SSRs) of *y* by distributed-lag stepwise multiple linear regression. Assume a time series matrix X whose columns, *x*, are standard or residual tree-ring chronologies. Each *x* is converted into a separate SSR of *y* by stepwise regression (Weisberg 1985) of *y* on *x* lagged -2 to +2 years from *y*. Such a regression, called "distributed lag" regression, has a long history of application in dendroclimatology (e.g., Stockton and Fritts 1973; Fritts 1976; Stockton and Meko 1983; Meko and Graybill 1995; Cook et al. 1999). The SSR model is cross-validated by leave-9-out cross-validation Meko (1997) and split-sample validation (Snee 1977) at each step as variables are entered in order of maximum reduction of the residual variance. Entry into the model is stopped if the next predictor to enter would result in decreased skill, as indicated by a drop in cross-validation RE (Fritts et al. 1990). This approach to stopping entry of predictors in stepwise regression is a cross-validation cutoff rule (Wilks 2019). Regardless of the change in cross-validation RE, entry of variable is not allowed to proceed beyond the step at which the model reaches maximum adjusted R^2 (Myers 1990). In summary, entry of variables into the SSR model is guided by a cross-validation cutoff rule and entry is limited by the maximum adjusted R^2.

TRISH can be run with either standard or residual chronologies (Cook and Kairiukstis 1990) as input. An option for whitening site chronologies within TRISH allows you to get a reasonable estimate of how residual chronologies would perform without actually uploading a matrix of residual chronologies to TRISH. From TRISH screens you can choose to prewhiten chronologies with an order-*p* autoregressive model (AR(*p*)) where *p* is 1 2 or 3. Prewhitening at the site chronology level (Meko et al. 1993) is not the same as prewhitening at the individual-core level, as done in computing residual chronologies, but the resulting chronologies are generally very similar.

Although the SSRs are specifically intended to deal with lagged response of tree-growth to climate that might vary site to site, it is possible to apply TRISH to reconstruction without lagged predictors. An option allows you to specify that lag-0 only is allowed in the pool of potential predictors.

The second stage of reconstruction is multi-site reconstruction, or MSR, whose objective is to combine the individual SSRs into a final reconstruction. In preparation for MSR, the SSRs are screened for strength and temporal stability of the signal for *y*, and the screened SSRs are converted to orthogonal variables by principal components analysis (PCA; Wilks 2019). The screened network includes only those SSRs with the following properties: 1) a statistically significant ($p<0.05$) overall-F of calibration, 2) positive skill of cross-validation, as reflected in a reduction-of-error (RE) statistic (Fritts et al. 1990) greater than zero, 3) positive skill of split-sample validation, as reflected in RE>0 for both halves of split-sample validation, and 4) a physically logical lag structure given the expected causative relationship of climate and tree rings. This last constraint means that the the SSR is rejected if the lags imply that the current year's *y* is predictable from just past years' *x*. Logically, tree-growth should not be able to respond to a climate fluctuation before that fluctuation occurs. You can make the screening less stringent by optionally removing the last two constraints (3 and 4) through input settings in TRISH.

As already mentoned, the SSRs are converted to orthogonal variables, or principal component (PC) scores. The PCA is done on the covariance matrix rather than the correlation matrix of the

screened SSRs. I prefer the covariance matrix because the individual variances of the SSRs are meaningful to the analysis – they reflect the strength of signal for *y*, which varies from chronology to chronology, or from one SSR to another. Namely, the variance of an SSR computed over its overlap with *y* is proportional to the variance of *y* explained by the SSR model.

The time series matrix of screened SSRs input to PCA covers all years in common to those SSRs plus any additional years until the last year of the most recently ending SSR. Any SSR with an earlier ending year than that last year is extended statistically by a quantile-analog procedure described later (see "**Figure03-SSR3**).

The last step in the MSR modeling is stepwise MLR on the PC scores of the screened SSRs. If the climatic (*y*) record is short and there are many PCs, the risk is high of a chance relationship, or overfitting of the regression model (e.g. Rencher and Pun 1980). A decision therefore has to be made on how many, and which, PCs to include in the pool of potential predictors for the MSR model. To lessen the chance of over-fitting, we restrict the size of the pool to a small fraction of the sample size (number of years) for calibration of the MSR model. The default in MLR1-PCA is to prioritize the PCs by their correlation with *y*, such that the pool of potential predictors includes only the *m* PCs most highly correlated with *y*. Specifically, by default, the pool is restricted to size $m<fN$, where f=0.10. In other words, by default, the pool of potential predictors is constrained to be less than 1/10 the number of observations in the calibration period (e.g., 100 years of calibration dictates no more than 9 PCs in the pool).

The user can optionally control the size of the pool of potential predictor from within TRISH in in one of two ways:

1. Use the correlation-based reduction just described, but specify some $f\neq0.10$.
2. Dispense with the correlation-based reduction, and instead specify that the the pool of PCs for the MSR include the *k* most important PCs of the SSRs as ranked by the eigenvalues.

To help with identificaton of *k* for option (2), a scree plot of eigenvalues against PC# is presented in a TRISH window. The user might observe, for example, that the first three PCs of the SSRs account for 90% of the variance of the SSRs, and decide to include just those PCs in the pool of potential predictors for the MSR model.

The MSR modeling itself is a done by forward stepwise regression of *y* on the PCs in the pool of potential predictors. A cross-validation stopping rule (e.g., Wilks 2019) is used to stop entry of predictors into the stepwise MSR model whenever the skill of prediction on independent data, as measured by cross-validation RE, would decrease with entry of an additional predictor. Regardless of the change in RE, the stepwise entry is not allowed to proceed beyond the step of "approximate maximum" adjusted R-squared of calibration (Myers 1990), meaning that adjusted R-squared at the next step increases by less than a specified increment. The increment is hard-coded in TRISH as 0.01. Therefore, for example, if at step 3 in stepwise modeling the adjusted R-squared is 0.500 and at step 4 the adjusted R-squared is 0.509, the stepwise entry would not be allowed to proceed beyond step 3.

Reconstruction in TRISH by method MLR1-PCA yields various files in an output directory:

1. **Sixteen figure files**. Figures 1-4 are for the SSRs. Figures 5-6 are for the PCA. Figures 7-16 are for the MSR. Files are described in more detail below.
2. **Eight tables of statistics**. Tables 1-2 are for the SSRs. Tables 3-4 are for the PCA (loadings; correlations of *y* with PCs. Tables 5-8 are for the MSR. Files are described in more detail below.
3. **Three files with tab-separated time series.** One file has the full-length PC scores (time series) of the screened SSRs. A second file has the calibration-period data of predictand and predictors

for the MSR model. A third file lists observed *y*, reconstructed *y*, and a 50% confidence interval for reconstructed *y*. Files are described in more detail below.

4. **Two miscellaneous files**. These files, MyDataMLR1.dat and progressTRISH.txt are used by the interactive TRISH tool, but are not important to the user and so are not described below.

FIGURE FILES (all .png)

1. **Figure01-SSR1. Summary of chronology screening by single-site reconstruction (SSR)**. Bar chart at left shows the number of chronologies at various stages of screening. Box plots at right summarize distributions of adjusted R-squared of SSR models for the subset of chronologies suitable for modeling. Screening tally as follows: 1) the N1 chronologies in the source tree-ring network, 2) the N2≤N1 chronologies in the user-drawn map polygon, 3) the N3 ≤N2 chronologies with sufficient time coverage for SSR modeling, and 4) the N4≤N3 chronologies passing screening for strength and temporal stability of hydrologic signal.

   Screening consists of several stages, starting with a user-specified network of N1 chronologies. First is the reduction to N2, which includes only those chronologies both in the user-drawn map polygon and with full time coverage over some specified interval. Second is reduction to N3, which includes only those chronologies whose time coverage has sufficient time coverage for some specified calibration interval. Third is reduction to N4, which includes only those chronologies passing SSR screening, described previously.

   Only N3 chronologies are SSR-modeled, and only N4 of those pass SSR screening. Because the F-level of adjusted R-squared is one of the screening criteria, the box plots usually indicate that the N4 screened chronologies have higher adjusted R-squared than the large set of N3 chronologies fit to SSR models.
2. **Figure02-SSR2 Bar charts summaring lags in SSR models**. At left is a histogram of the number of models that include lags t-2 to t+2 relative to *y* for the N3 chronologies subjected to SSR modeling. At right is a similar histogram for just the N4 chronologies passing screening.

   The histograms are useful for a quick assessment of whether lags are important for the chronologies modeled. It is not unusual for some models to include positive lags but not lag-0, meaning a delayed response of growth to climate.
3. **Figure03-SSR3**. **Drop in maximum SSR signal-strength with loss of chronologies toward present**. Depending on the input specifications, SSR models may have been fit to a specific uniform calibration period, or to whatever overlap period of y is available for each chronology (variable calibration period). The maximum R-squared (left y-axis) applies to that calibration period. Chronologies generally have different ending years, so that the nearer to the present the fewer the available chronologies (or SSRs), because each chronology is modeled by an SSR).

   The red line shows the drop in available chronologies from an initial year when all N4 SSRs are available, to the last year with any chronologies are available. The black line, and right y-axis, shows the corresponding maximum adjusted R-squared of the SSR models that extend to the given year. For example, if the SSR with the strongest signal (highest R-squared) also happens to be the chronology with the most recent time coverage, the the black line will be at a constant over the time period of dropping sample size.

   The plot is intended as interactive feedback to allow you to pick a reasonable ending year for the calibration period of SSR models, and equivalently, of the final MLR model. You would not want the calibration period to extend past the last year of tree-ring data with a strong SSR model. For example, if all screened SSRs were complete through 1998, and if chronologies dropped out such that only one was available in 2017, you could specify an ending year for

calibrating the MSR model anywhere in the interval 1998-2017. If the chronology with the strongest signal (highest R-squared of SSR) had the end year 2017, it might make sense to to specify the end year for MSR calibration as 2017. But if there was a big drop in maximum R-squared after some year (say, 2004), you might want to specify 2004 as the end year for calibration of the MSR.

The currently specified last year of the default calibration period is annotated on the figure as "calibration end year." Calibration of models is not allowed to include any *y* data after that year. The actual calibration periods of SSR models might not extend to the specified year, depending on the ending date of chronologies. "NA" as "calibration end year" means that you have allowed the data itself to set the maximum possible calibration period depending on the overlap of the chronology with *y*. If you happen to set the calibration end year to "NA," the actual calibration period for all SSR models will extend to the year of the most recently collected tree-ring chronology. Calibration periods of individual SSR models are listed in the tables in output files Table1-SSR1.txt and Table2_SSR1.txt.

How is it possible to specify and ending year of the calibration period later than the end of one of the chronologies. This is possible because before MLR modeling, TRISH extends all of the SSRs by a quantile analog method to the year of the most recently ending SSR. Extension is done by combination of correlation analysis and quantile ranking, as follows, assuming **A** is the SSR needing extended in year *k*. (1) A common period of length at least 50 years is identified for which all SSRs are complete – no missing data. (2) The correlation matrix (Spearman) of SSRs is computed for the common period, and correlations are sorted from largest to smallest. (3) The SSR that is most highly correlated with **A**, has data in year *k*, and has at least 100 year of overlap of complete data with **A** is defined as the predictor SSR, or series **B**. (4) The quantile of the value of **B** in year *k* is computed in the overlap period of **A** and **B**. (5) The same quantile of **A** in that overlap period is defined as the estimate of **A** in year *k*. Steps 3-5 are repeated until all SSR series have been extended to the ending year of the most recently ending SSR.

To avoid using extended SSRs in the reconstruction, specify at the TRISH input window an ending year of calibration that is the same as the last year of coverage by all of the screened SSRs before extension. This year can be identified from Figure 3.

4. **Figure04-SSR4**. **Scatterplot of the mean of the N4 SSRs of *y* against observed y**. The plot also has a fitted least-squares straight line, a locally-weighted (loess) fit to the scatter of points, and separate loess fits to the positive and negative residuals from the loess-fit to the points. The correlation coefficient is annotated on the plot. This figure plot gives a quick view of the strength of signal for *y* in the mean of the screened SSRs and whether the relationship of observed *y* with the mean of the SSRs is linear, curvilinear, restricted to one particular part of the range of *y*, or driven by outliers.
5. **Figure05-PCA1**. **Scree plot of eigenvalues of PCs on the screened subset of SSRs**. The percentage of variance and cumulative percentage of variance of SSRs accounted for by the first few (up to 7) PCs is annotated on the plot. A horizontal dashed line marks the "average" eigenvalue. For uses wanting to specify how many of the "most important" PCs to retain for the pool of potential predictor for the MSR model, those PCs above the dashed line are consistent with an "eigenvalue of 1" cutoff if the PCA had been done on the correlation matrix (It was done on the covariance matrix.)
6. **Figure06-PCA2**. **Heat map of of PC loadings**. Heat map includes all PCs of the screened SSRs. The heat map summarizes the modes of variability of the tree-ring chronologies after they have been converted to SSRs by lagged regression. Each column gives the loadings of a

PC on the individual chronologies. These loading are also listed in one of the output tables. Loading of all the same sign indicate a mode of variability in common to to all chronologies. Loading of opposite sign indicate contrasts in the component of variability of the SSR predictand, *y*, contributed by various chronologies. A contrast might be driven by regional contrast in climate signal (some chronologies in the north, others in the south, and some years having contrasting N/S moisture anomalies). Another possible reason for a contrast is some chronologies responding primarily to fall precipitation and others to winter precipitation: if *y* is water-year discharge, say, of a river, different subsets of chronologies could contribute to the signal for *y*, but the contributions are not necessarily positively correlated.

7. **Figure07-Calibration1**. **Bar plot of correlations and autocorrelations**. Blue bars are Pearson correlation of the predictand, *y*, with the PCs of the screened SSRs. Magenta bars are lag-1 autocorrelation of the PCs. A 95% confidence band on the correlations is plotted by default. (User can change this to 90% or 99% within TRISH.) The statistical significance of the computed correlations of *y* with PCs is not adjusted for autocorrelation in the time series. The autocorrelation of *y* is shown with a horizontal magenta line. Difference in the height of this line and the heights of the magenta bars shows whether the PC scores of the SSRs are more or less autocorrelated than the hydrologic variable to be reconstructed. It is important to recognize that the lag-1 autocorrelation can reflect long-memory (e.g., trend) as well as short-memory processes.

   The correlation plotted as a blue bar directly gives the strength of linear correlation of *y* with a specific PC of the SSRs. Futher interpretation in terms of direction of influence of individual chronologies is complicated, and depends on several factors, including the sign of the PC loading of the SSR, and the signs of the regression coefficients in the lagged regression model that converts the chronology to the SSR. A more direct way of assessing importance of individual chronologies to *y* is from the coefficients and statistics of the SSR models (**Table1-SSR1.txt**, see below)

   The correlation bars in Figure 7 are also useful in that they display the priority for inclusion of PCs in the pool of potential predictors for the MSR model. Priority is in order of absolute correlation with *y.* However, all PCs are allowed in the pool as long the number of available PCs is less than some specified fraction (default 1/10) of the number of years in the calibration period (see above).

8. **Figure08-Calibration2**. **Summary of calibration of MSR model**. The scatterplot of reconstructed *y* against observed *y* graphically shows the accuracy of the reconstruction. Curvature in the scatter could indicate non-linearity poorly handled by the model. Fanning out of points in the scatter toward the right, say, indicates heteroskedasticity of errors, in that error variance increases toward higher observed *y*.

   The time series plots of reconstructed and observed y show the ability of the reconstruction to track temporal variation of *y* in the calibration period. Years or multi-year period of good or bad agreement are readily seen in the plots. Any diverging trends in observed and reconstructed y can also be seen in these plots. The plots in this way can visually corroborate any significant trend in residuals that might be identified in the analysis of residuals described later.

   The annotated MSR calibration statistics summarize the final regression model. Statistics generally in tree-ring reconstruction papers are included. Statistics include the size of the pool of potential predictors and the number of those included in the final model. More detailed information on the model can be found in output Table 5 and Table 8 (described below).

9. **Figure09-Calibration3**. **Histograms and autocorrelation functions (acf's) of observed and reconstructed predictand for the calibration period**. The histograms at left allow comparison of distributions of observed and reconstructed *y*. These histograms have the same x-axis scale so that the compression of variance imposed by regression is readily seen: the distribution of reconstructed *y* is compressed relative to that of observed y. Such compression is a property of regression, and reflects the inability of the tree-ring data to to explain all of the variance of *y*. This compression is not a flaw in the regression process. It merely reflects that tree rings are imperfect recorders of *y*. The temptation to "scale" the variance of the reconstruction by amplifying the departures from the mean such that the variances of reconstructed and observed *y* are equal in the calibration period should be resisted because the sign of the noise component for years before the calibration period is unknown (residuals assumed to have zero mean and constant variance).

   The acf's allow comparison of the autocorrelation properties of observed and reconstructed *y* over a common period. The acf reflects persistence of above-mean or below-mean values, and so is related to duration of wet and dry periods. Greatly different acf's could indicate that the reconstruction overestimates or underestimates such duration.
10. **Figure10-AnalysisResiduals1. Analysis of residuals for normality and constancy of variance**. Analysis of residuals is an essential part of any regression modeling, and has the goal of checking that the regression assumptions on statistical properties of residuals are not violated. Texts on regression (e.g., Draper and Smith 1981; Myers 1990) cover most aspects of analysis of residuals relevant to TRISH.

    Normality can be checked visually by the histogram of residuals, at left, and statistically by the annotated results of the Lilliefors Test (Conover 1980). The *p*-value for Lilliefors Test, annotated at top of the histogram, indicates whether we can reject the null hypthesis that the residuals come from a normal distribution.

    Constancy of variance is checked visually with the scatterplot of residuals against predicted values (right), and statistically tested with the Breusch-Pagan (Breusch and Pagan 1979). The scatter plot ideally shows no noticeable pattern (e.g., fanning out, curvature) that might be evidence of violation of the assumption of constancy, or homogeneity, of variance of residuals (Myers 1990). The *p*-value for the Breusch-Pagan test should be greater than 0.05. Otherwise you must reject the null hypothesis of homogeneous variance of residuals, and conclude that the residuals are "heteroskedastic." Heteroskedastic means that there is a dependence of variance of residuals on the fitted values of the regression. Visually, one common example of heteroskedasticity residuals that fan out, or become more spread, toward higher fitted values.
11. **Figure11-AnalysisResiduals2**. **Analysis of residuals for trend.** Regression residuals are plotted as a time series and the the time plot is tested for trend using the non-parametric Mann-Kendall trend test (Wilks 2019). The plot includes a non-parametric best-fit line,following Haan (2002), and is annotated with results of the Mann-Kendall test and with the estimates of the fitted parameters of the line as well as a variance inflation factor (VIF). The VIF reflects possible adjustment of the significance of trend for autocorrelation in the residuals – that is, autocorrelation in addition to that due to trend. The autocorrelation adjustment follows Wilks (2019), and its application is indicated by a VIF>1.0. If VIF=1.0, the residuals after removal of linear trend have no significant lag-1 positive autocorrelation (one-sided test, α=0.05).

    Residuals ideally have no trend. Trend might indicate that some factor, with trend, other than the reconstruction predictand, is influencing the tree-ring index. There could be other causes (e.g., nonclimatic growth trend still in tree-ring index after standardization).

12. **Figure12-AnalysisResiduals3. Analysis of residuals for autocorrelation.** The acf of residuals is plotted with a 95% confidence interval for lags k=0 to k=m years, where m is the minimum of 20 and N/4, where N is the length of the time series of residuals. The acf is annotated with results of a Durbin-Watson (DW) test of the null hypothesis that the population lag-1 autocorrelation of regression residuals is zero (Myers 1990). Ideally, the DW statistic is not significant, and the acf is within its 95% confidence interval for all lags k>0.

    One of the regression assumptions is that the residuals have zero autocorrelation. Autocorrelated residuals could mean that the persistence of droughts and wet periods, say, is poorly represented in the reconstruction. Use of lags in the regression model, or use of residual rather than standard chronologies, can affect the autocorrelation of the regression residuals.
13. **Figure13-Validation1**. **Cross-validation predictions**. The observed y is plotted as a time series along with two versions of predicted y: 1) as provided by the reconstruction model, and 2) as provided by cross-validation. The latter are the cross-validation predictions. The two versions of predicted y will differ more and more as validation becomes poorer. Note that the cross-validation is leave-m-out rather than leave-1-out to ensure that with a lagged regression model, tree-ring data used to provide a cross-validation prediction are not also used to calibrate the model that gives that prediction (Meko 1997). How many observations are left out in cross-validation is computed as m=1+4k, where k is the maximum positive or negative lag allowed on the predictors in the SSR regression models. For example, if lags -2 to + 2 years are included in the pool of potential predictors, m=1+4(2)=9 observations left out. If no lags are used (k=0), leave-m-out cross-validation reduces to leave-1-out cross-validation.
14. **Figure14-Validation2. Validation summary statistics**. This figure summarizes the cross-validation and split-sample validation of the MSR model. At left is a histogram of the cross-validation residuals, with annotated test results (Lilliefors test) of the null hypothesis that those residuals are from a normal distribution of unspecified mean and variance. Ideally, these residuals are normally distributed. A test for normality is also applied to the regression residuals (Fig 7), and because those residuals generally are similar to the cross-validation residuals, the Lilliefors test result are also generally similar.

    At right are annotated statistics for cross-validation and split-sample validation of the MSR model. The root-mean-square error (RMSE) of cross-validation is especially useful because combined with an assumption of normality, the RMSE of cross-validation can be applied to place confidence bands on the annual reconstructed y. For example, a 95% confidence band is yhat±1.96 RMSEcv, where yhat is the reconstructed y, and RMSEcv is the cross-validation RMSE.

    The reduction-of-error (RE) statistic is a measure of skill of the reconstruction when applied to data not used in calibrating the reconstruction model. RE is reported here for both cross-validation and split-sample validation (Fritts et al. 1990). The split-sample validation by default uses an even split on the overlap of y and x, with the early split one observation longer than the later split when the full overlap is an odd number of years. The statistic RE is not assigned "significance." The typical interpretation is that RE>0 is required for the model to have ANY skill relative to simply substituting the calibration-period mean of y as the reconstruction for each year.
15. **Figure15-Reconstruction1**. **Full-length reconstruction with confidence interval**. The full-length reconstruction of y is plotted with its 50% confidence interval computed as yhat±0.67449 RMSEcv where yhat is reconstructed annual y. Note that 0.6449 is the 0.75 quantile of the standard normal distribution and RMSEcv is the root-mean-square error of cross-validation of the MSR model. This confidence interval therefore assumes normally distributed errors.

The last few years of the reconstruction will have amplified uncertainty if based on statistically extended SSRs (see description of Fig. 3, above). For that reason, you may want to truncate the reconstruction to end with the last year that all of the screened SSRs (before extension) have data. That way all years of the reconstruction are base on observed tree-ring chronologies rather than possibly being distorted because some tree-ring chronologies (their SSRs) are extended from other chronologies.

16. **Figure16-Reconstruction2**. **Autocorrelation functions (acf) and boxplots of reconstructed y for the calibration period of the MSR model and for earlier years**. These plots allow a quick graphical assessment of the calibration period *y* in a long-term perspective. The comparisons here are "apples vs apples" in that reconstructed *y* is compared with reconstructed *y* for different time intervals. It is of course possible to compare the pre-calibration reconstruction with observed *y* for the calibration period, but here variance of the reconstruction is necessarily compressed relative to observed *y* by the regression process. Such a comparison might incorrectly suggest that variance of *y* has increased in modern times.

TABLES OF STATISTICS (all tab-sep .txt)

When viewing tables on screen with text editor, use monospaced font; otherwise columns will not line up properly with headings.

1. **Table1-SSR1.txt.** Summary statistics of SSR models for all chronologies fit with reconstruction models.
    1. **N_1**: sequential number in this table
    2. **N_2**: corresponding site number in original tree-ring network
    3. S**ite**: unique alphanumeric identifier for tree-ring chronology
    4. **Goc**: first year of model calibration period
    5. **Endc**: last year of model calibration period
    6. **Model**: code indicating lags in model and order that they entered stepwise. The five slots correspond to lags -2, -1, 0, +1, +2 years relative to the predictand year. For example, code [0 2 1 0 0] means lag 0 entered first, lag t-1 entered second, and no other lags are in the model.
    7. **Sign**: code that goes along with "Model" and tells the sign of the SSR regression coefficients on the the chronology lagged -2, -1, 0, +1, +2 years relative to the predictand year. Codes "P", "N" ad "0" indicate positive coefficient, negative coefficient, and not in model. For example, [00P0P] means positive coefficent on lag 0, and positive coefficient on lag t+2.
    8. **R2a**: adjusted R-squared of model; this is the regression R-squared adjusted downward as a penalty for number of predictors in the model.
    9. **pF**: p-value of the overall F of regression (pF<0.05 indicates significant model at 0.05 level)
    10. **REcv**: Reduction of error (RE) statistic from leave-9-out cross-validation
    11. **REa**: Split sample RE for fitting model to first half of record and validating on second half
    12. **REb**: Split sample RE for fitting model to second half of record and validating on first half
    13. **Refit**: Logical (TRUE or FALSE) variable indicating whether the model was re-fit with expanded calibration period after exploratory stepwise regression.

        First the stepwise regression allows lags -2 to +2 in the model, which could restrict the calibration period if the tree-ring data happen to end in the same year or in the year after the end of y. For example, if y and the chronology end in 2019, the calibration period cannot have an ending year later than 2017, because of the possible need for lags +1 and +2 on the tree-ring series. The stepwise process might result in a model that does not include lags +1 or +2. If so, the model is re-fit, resulting in the maximum possible length of calibration period given the lags in the model and time coverage of y and the tree-ring chronology. "Refit" indicates whether the model was re-fit. .
    14. **Gor**: First year of reconstruction.
    15. **Endr**. Last year of reconstruction.
    16. **Reject**: Logical (TRUE or FALSE) variable indicating whether the chronology and its SSR are rejected from further use in the later step of multi-site reconstruction. Rejection occurs if any of the following are true:
        1. pF≥0.05
        2. REcv≤0
        3. REa≤0 or REb≤0

4. Illogical causal model: the final model has y predicted from x at negative lags only from y. This is physically unreasonable, because past years' tree-ring values alone should not be able to detect current year's climate.

2. **Table2-SSR2.txt.** Summary statistics of SSR models for just those chronologies passing screening for hydrologic signal. These are the chronologies for which "Reject" is FALSE in Table 1. The columns are the same as those of Table 1.
3. **Table3-PCA1.txt.** Loadings of PCs of the screened subset of SSRs.
   These loadings are the PCA weights of each PC on the SSRs derived from individual tree-ring chronologies. Each SSR is associated with a specific tree-ring chronology, as indcated by the column "SiteID." The site number, in the user database supplied to TRISH, is listed in column "Site#." The percentage of variance of SSRs accounted for by the PCs applies to the full overlap of the SSRs, not just the period in common with the hydrologic variable. Note that the PCA was done on that full period of overlap. Accordingly, PC scores for the full overlap are by design not intercorrelated, but they may be intercorrelated during the shorter overlap with the hydrologic variable.
4. **Table4-PCA2.txt.** Correlation of y with PCs of the screened subset of SSRs.
   The analysis is done on the calibration period of the MSR model. Thresholds (95%) are shown for significance disregarding (Thresh1) and considering (Thresh2) lag-1 autocorrelation in *y* and the PCs. The sign of correlation of *y* with a PC relates to the direction of influence of high or low index for a particular chronology to reconstructed *y,* but the interpretation is complicated. Interpretation must consider 1) he signs of the lagged regression coefficients of the SSR model that converts the chronology to a SSR, 2) the signs of loadings of the PC on the SSRs (Figure 6), and the signs of the coefficients of the MSR model that converts the PC scores to a time series of reconstructed *y* (Table 8).
5. **Table5-Calibration1.txt.** Calibration statistics of the multi-site reconstruction (MSR) model.
   1. **YearGo:** first year of calibration peirod
   2. **YearStop:** last year of calibration period
   3. **Npool**: number of PCs in the pool of potential predictors
   4. **Npredictors**: number of predictors (PCs) in MSR model
   5. **R2:** regression R-squared
   6. **F:** overall-F of regression
   7. **pF:** *p*-value of F
   8. **R2adj**: adjusted regression R-squared
   9. **RMSEc** is root-mean-square error of calibraton, which is also called the "standard error of the estimate.
6. **Table6-AnalysisResiduals1.txt**. Normality, autocorrelation, trend, heteroskedasticity
   1. **YearGo:** Start ear of calibration period
   2. **YearStop**: End ear of calibration period
   3. **pNormal**: p-value of Lilliefors test for normality (pNormal<0.05: reject H0 that residuals from normal distribution)
   4. **DW:** Durbin-Watson statistic
   5. **pDW**: Durbin-Watson statistic p-value (pDW<0.05: reject H0 that population lag-1 autocorrelation of regression residuals zero)
   6. **TrendSlope**: slope coefficient of non-parametricfit of trend of residuals
   7. **pTrend:** Slope of trend line, with p-value from Mann-Kendall test
   8. **BP Test ChiSq:** Breusch-Pagan Chi-squared statistic (test for homoskedasticity of residuals)
   9. **dfBP:** degress of freedom for BP test: equals number of predictors in model 1

10. **pBP:** *p*-value for BP test; pBP<0.05 indicate reject H0 that residuals homoskedastic at 0.05 level

7. **Table7-Validation1.txt**. Cross-validation and split-sample validation statistics of MSR model.
    1. **NleaveOut**: Cross-validation how many left out
    2. **RMSEcv**: Cross-validation root-mean-square error
    3. **REcv**: reduction of error (RE) statistics from cross-validation
    4. **YearGoA**, **YearStopA**: first and last years of ealy split-sample period
    5. **YearGoB, YearStopB**: first and last years of ealy split-sample period
    6. **REsplitA**: split-sample RE for calibration on early (A) and validation on late (B)
    7. **REsplitB**: split-sample RE for calibration on late (B) and validation on early (A)
8. **Table8-CoefficientsMSR.** Estimated regression coefficients of the MSR model
    1. **Intercept:** constant term in the regression
    2. <remaining rows>: each row corresponds to a PC included in the model; variables are labeled by the PC number (e.g., “PC4”). The signs of the coefficients are related to whether tree rings were wide or narrow with higher or lower values of the hydrologic variable, but interpretation must include a look at the PC loading on the tree-ring chronologies. Those loadings are heat-mapped in Figure 6.

TIME SERIES OUTPUT

1. **PCscoresTimeSeries.** Scores of all PCs of the screened SSRs. All of these PCs may not actually enter the final reconstruction model. Note that the subset of scores for just those PCs in the MSR model, for the calibration period, are included after the predictand column in "RegressionInputTimeSeries.txt." First, the restriction of the pool of potential predictors could have been reduced because it is restricted to be smaller than some decimal fraction (input f) times the number of years in the calibration period. Second, the stepwise regression might select just some of the predictors from the pool of potential predictors.
2. **RegressionInputTimeSeries.txt.** Listing of predictand and predictor time series for the calibration period of the MSR model. User could repeat the regression model outside of TRISH by regressing the data in the second column on the data in the remaining columns.
   1. **Year:** year of data
   2. <*predictand (units)*> The predictand for the regression model
   3. <*predictors – PCs*>: the scores of the PCs in the MSR model, in same order as the coefficients are listed in Table 8.
3. **ReconstructedWithConfidenceIntervalTimeSeries.txt**. Listing of time series of observed predictand, reconstruction, and confidence interval on reconstruction (5 columns). Call the observed predictand *y* and the reconstruction yhat. Confidence interval is estimated assuming that reconstuction errors are normally distributed with a standard deviation equal to RMSEcv.
   1. **Year:** year of data
   2. **y**: observed predictand
   3. **yhat:** the reconstruction
   4. **Lower 50% CI:** true (unknown) value of predictand has 50% chance of being lower than this threshold of the confidence interval (CI)
   5. **Upper 50% CI:** true (unknown) value of predictand has 50% chance of being higher than this threshold of the confidence interval (CI)

REFERENCES

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroskedasticity and random coefficient variation. Econometrica, 47 (5), 1287–1294.

Conover, W. (1980). Practical nonparametric statistics (Second ed.). New York: John Wiley & Sons. (493 pp)

Cook, E. R., & Kairiukstis, L. A. (Eds.). (1990). *Methods of dendrochronology: Applications in the environmental sciences*. Dordrecht: Kluwer Academic Publishers. (394 pp)

Cook, E. R., Meko, D. M., Stahle, D. W., & Cleaveland, M. K. (1999). Drought reconstructions for the continental United States. J. Climate, 12 , 1145-1162.

Draper, N. R., & Smith, H. (1981). Applied regression analysis. New York: John Wiley Sons. (709 pp)

Fritts, H. C. (1976). Tree rings and climate. London: Academic Press. (567 pp)

Fritts, H. C., Guiot, J., & Gordon, G. A. (1990). Verification. In E. R. Cook & L. A. Kairiukstis (Eds.), Methods of dendrochronology: Applications in the environmental sciences (pp. 178–185). Kluwer Academic Publishers.

Haan, C. T. (2002). Statistical methods in hydrology (Second ed.). Iowa State University Press. (496 pp)

Hutton, P. H., Meko, D. M., & Roy, S. B. (2021). Supporting restoration decisions through integration of tree-ring and modeling data: reconstructing flow and salinity in the San Francisco Estuary over the past millennium. Water , 13 (15), 2139. ('Received 2 June 2021, revised: 21 July 2021, accepted 27 July 2021, published 3 August 2021; this article belongs to the Special Issue Decision Support Tools for Water Quality Management)) doi: 10.3390/w13152139

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

Meko, D. M., Cook, E. R., Stahle, D. W., Stockton, C. W., & Hughes, M. K. (1993). Spatial patterns of tree-growth anomalies in the United States and southeastern Canada. J. Climate, 6 , 1773–1786.

Meko, D., & Graybill, D. A. (1995). Tree-ring reconstruction of Upper Gila River discharge. J. Am. Water Resour. Assoc., 31 (4), 605–616.

Meko, D. M., Woodhouse, C. A., Baisan, C. H., Knight, T., Lukas, J. J., Hughes, M. K., & Salzer, M. W. (2007). Medieval drought in the Upper Colorado River Basin. Geophys. Res. Lett., 34 (L10705), 10.1029/2007GL029988.

Myers, R. H. (1990). Classical and modern regression with applications. Pacific Grove, California: Doxbury. (488 pp)

Stockton, C. W., & Fritts, H. C. (1973). Long-term reconstruction of water level changes for Lake Athabasca by analysis of tree rings. J. Am. Water Resour. Assoc., 9 (5), 1006–1027. doi: 10.1111/j.1752-1688.1973.tb05826.x

Stockton, C. W., & Meko, D. M. (1983). Drought recurrence in the Great Plains as reconstructed from long-term tree-ring records. J. Clim. Appl. Meteor., 22 , 17-29.

Weisberg, S. (1985). Applied linear regression, 2nd ed. New York: John Wiley. (324 pp)

Wilks, D. S. (2019). Statistical methods in the atmospheric sciences (Fourth ed.). Cambridge, MA: Elsevier. (818 pp)

Woodhouse, C. A., Meko, D. M., & Bigio, E. R. (2020). A long view of Southern California water supply: perfect droughts revisited. J. Am. Water Resour. Assoc., 56 (2), 212–229. doi: https://doi.org/10.1111/1752-1688.12822