

Monthly Paleostreamflow Reconstruction from Annual Tree-Ring Chronologies

J.H. Stagge^{a,*}, D.E. Rosenberg^a, R.J. DeRose^b, T.M. Rittenour^c

^a*Utah State University, Department of Civil and Environmental Engineering, Logan, UT 84321*

^b*United States Forest Service, Forest Inventory and Analysis, Rocky Mountain Research Station, Ogden, UT 84401*

^c*Utah State University, Department of Geology, Logan, UT 84321*

Abstract

Paleoclimate reconstructions are increasingly used to characterize annual climate variability prior to the instrumental record, to improve estimates of climate extremes, and to provide a baseline for climate-change projections. To date, paleoclimate records have seen limited engineering use to estimate hydrologic risks because water systems models and managers usually require streamflow input at the monthly scale. This study explores the hypothesis that monthly streamflows can be adequately modeled by statistically decomposing annual flow reconstructions. To test this hypothesis, a model for monthly streamflow reconstruction is presented that uses multiple linear regression, but expands the set of predictors to include annual streamflow reconstructions, reconstructions of global circulation, and potential differences among regional tree-ring chronologies related to tree species and geographic location. This approach is used to reconstruct 600 years of monthly streamflows at two sites on the Bear and Logan rivers in northern Utah. Nash-Sutcliffe Efficiencies remain above zero (0.26-0.60) for all months except April and Pearson's correlation coefficients (R) are 0.94 and 0.88 for the Bear and Logan rivers, respectively, confirming that the model can adequately reproduce monthly flows during the reference period (10/1942 to 9/2015). Incorporating a flexible transition between the previous and concurrent annual reconstructed flows was the most important factor for model skill. Expanding the model to include global climate indices (ENSO and PDO) and regional tree-ring chronologies produced smaller, but still significant improvements in model fit. The model presented here is the only approach currently available to reconstruct monthly streamflows directly from tree-ring chronologies and climate reconstructions. With reasonable estimates of monthly flow that extend back in time many centuries, water managers can challenge systems models with a larger range of natural variability in drought and pluvial events and better evaluate extreme events with recurrence in-

tervals longer than the observed record. Establishing this baseline is critical when estimating future hydrologic risks under conditions of a non-stationary climate.

Keywords: paleoclimate, streamflow reconstruction, tree-ring, global circulation, dendrochronology

1 1. Introduction

2 While grappling with the question of how future climate changes could affect the likelihood
3 and severity of hydrological extremes (floods and droughts), hydrologists, engineers, and water
4 resources planners have acknowledged the potential for streamflow reconstructions to charac-
5 terize pre-industrial hydrologic variability over multiple centuries (Bonin and Burn, 2005). By
6 combining reconstructions of the past with climate change projections, it may be possible to place
7 the signal of climate change-induced streamflow trends in the context of natural variability. In
8 addition, streamflow reconstructions can significantly increase the number of scenarios used for
9 water resources simulation or optimization. Despite these potential benefits, streamflow recon-
10 structions have not gained widespread use in water systems analysis in part because flow has been
11 typically reconstructed at an annual resolution, which is generally too coarse for analysis of water
12 vulnerability and decision-making. This study explores the hypothesis that monthly streamflows
13 can be adequately predicted from available data and presents a novel method to perform such
14 a reconstruction based on temporal disaggregation of the original annual flow reconstruction,
15 spatial and species-based differences in regional tree-ring chronologies, and reconstructions of
16 global circulation indices.

17 Existing techniques for streamflow reconstruction primarily rely on linear regression to re-
18 late carefully chosen tree-ring chronologies to mean annual flow (MAF) during the available
19 instrumental record, and then extending the historical timeseries by applying this relationship
20 to the full tree-ring record (Loaiciga et al., 1993). This approach assumes that the instrumental
21 record is representative of the reconstructed period and that changes in streamflow are driven
22 solely by changes in climate. As a result, this approach works best when applied to river reaches
23 without significant man-made impoundments or where those effects have been removed to gen-

*Corresponding author
Email address: james.stagge@usu.edu (J.H. Stagge)

24 erate a near-natural observed record. Annual streamflow reconstructions are commonly available
25 and have been developed globally (Meko et al., 2001; Woodhouse and Lukas, 2006; Woodhouse
26 et al., 2006), wherever sufficient tree-ring and streamflow data are available.

27 Despite the availability of annual streamflow reconstructions, only one methodology cur-
28 rently exists to reconstruct monthly flow from annual resolved tree-ring chronologies (Gan-
29 gopadhyay et al., 2015). This method resamples annual subsets of temperature and precipitation
30 from the instrumental records and matches them to tree-ring widths in the paleo-record using a
31 K-nearest neighbor approach repeated many times to develop an ensemble of temperature and
32 precipitation timeseries (Gangopadhyay et al., 2009). Temperature and precipitation are then
33 used as inputs for a water balance model (Wolock and McCabe, 1999). This method has been
34 used to generate seasonal streamflow in Nevada (Solander et al., 2010) and monthly stream-
35 flow along the Colorado River (Gangopadhyay et al., 2015). While useful for generating runoff
36 in well-studied watersheds, this approach is somewhat limited because a calibrated watershed
37 model is not always available. Additionally, by resampling from the observed record, the po-
38 tential monthly time series of temperature and precipitation are limited to re-ordering \approx 60-100
39 observed annual subsets from the instrumental record. While this may be an effective approach
40 for some locations, alternative methods that can instead reconstruct monthly streamflow directly,
41 without the need for watershed models and the limitations of repeated resampling, are highly
42 desirable.

43 This study introduces and tests a novel approach for reconstructing monthly streamflow di-
44 rectly from available tree-ring chronologies and climate reconstructions. We outline a model-
45 ing framework that builds in complexity and number of predictors by considering: 1) annual
46 streamflow reconstructions; 2) regionally available tree-ring chronologies, and; 3) global cli-
47 mate reconstructions of El Niño-Southern Oscillation (ENSO) and Pacific Decadal Oscillation
48 (PDO). To demonstrate this flexible approach, model fits and predictive skills are determined for
49 two streamflow gages in northern Utah where annual streamflow reconstructions have previously
50 been developed. Model results are further evaluated and discussed in the context of individual
51 predictors, their physical basis, and implications for water management.

52 **2. Models**

53 Three model frameworks are introduced in this study as potential candidates for the recon-
54 struction of mean monthly streamflow. The Monthly Fraction (MF) model uses simple assump-
55 tions to reconstruct monthly streamflows and is included as a null model, against which the other
56 models can be compared. The remaining two models constitute the primary approach proposed
57 herein and are presented in order of increasing complexity, each applying the same basic frame-
58 work, but with additional predictors. First, the Annual Percentile (AP) model directly links the
59 reconstructed annual streamflow percentiles to monthly percentiles. Second, the Annual Per-
60 centile with Regression (APR) model instead estimates monthly percentiles using regression,
61 first considering only lagged annual streamflow percentiles and ultimately incorporating other
62 predictors such as global climate indices or spatial/species-specific patterns in regional tree-ring
63 chronologies. All code and data for these models is freely available in an online repository at
64 https://github.com/jstagge/monthly_paleo.

65 *2.1. Monthly Fraction (MF) Model*

The Monthly Fraction (MF) model assumes that the monthly proportion of mean annual streamflow is identical across all years. Based on this assumption, the monthly fraction, $\bar{f}_{m,y}$, of the MAF is determined for each month by dividing the monthly flow volume, $Q_{m,y}$, by annual volume, MAF_y . In this notation, “m” and “y” subscripts correspond to monthly and annual steps, respectively. Monthly streamflow is then reconstructed by multiplying the reconstructed MAF by the appropriate mean monthly fraction, \bar{f}_m , calculated for each of the 12 months :

$$\hat{Q}_m = \bar{f}_m \times MAF \times 12 \quad , \text{where} \quad \bar{f}_m = \frac{\sum_{y=1}^n f_{m,y}}{n} \quad \text{and} \quad f_{m,y} = \frac{Q_{m,y}}{MAF_y} \quad (1)$$

66 where \hat{Q}_m represents estimated monthly streamflow. Monthly flows reconstructed by the MF
67 model retain the same seasonal shape, but are scaled linearly. For example, if an average of 30%
68 of each years historical flow volume occurred during June, this proportion is maintained in the
69 reconstruction. Prior to performing reconstruction, the 12 F_m values were re-scaled to sum to 1,
70 ensuring that the MAF would be maintained. The mean F_m was used as the monthly fraction,
71 rather than the median, which made little difference in the case study results.

72 2.2. Annual Percentile (AP) Model

73 The Annual Percentile (AP) model forms the basis for all subsequent reconstruction models.
74 It assumes that the streamflow percentile for each month is equivalent to the MAF percentile
75 for the particular water-year. That is, if a given years MAF is in the 20th percentile, then the
76 mean flow for each month of that water-year was also the 20th percentile, determined by the
77 cumulative probability distribution of flow for each month. Because subsequent models depend
78 on a Gaussian distribution, we transformed the percentile (0-100) to a standard normal value
79 ($-\infty - \infty$), represented by the variable N :

$$\hat{N}_m = N_a \quad (2)$$

80 where \hat{N}_m is the estimated monthly standard normal and the subscripts m and a again represent
81 monthly and annual values of the normalized streamflow percentile.

82 For use in the AP model, a univariate probability distribution was fit to the MAF and each
83 of the monthly mean flows to permit transformation to and from the standard normal distribu-
84 tion. Non-parametric alternatives, such as empirical cumulative distributions, were not consid-
85 ered because they limit the ability to extrapolate beyond the original data. Annual and monthly
86 cumulative distributions were fit based on Maximum Likelihood Estimation (MLE) using the
87 fitdistrplus package in R (Delignette-Muller and Dutang, 2015). Candidate distributions were vi-
88 sually compared using skewness-kurtosis plots (Cullen and Frey, 1999) evaluated with 500 non-
89 parametric bootstrap realizations. Final selection of candidate distributions was further validated
90 based on Akaike information criterion (AIC), quantile plots, and results of Kolmogorov-Smirnov
91 and Anderson-Darling tests. Statistical distribution tests were based on 5,000 bootstrap simula-
92 tions, thereby avoiding the issue of applying tests where parameters are estimated from the data
93 (Crutcher, 1975).

94 2.3. Annual Percentile with Regression (APR) Model

95 The Annual Percentile with Regression (APR) model has the same fundamental structure as
96 the AP model without enforcing a 1:1 relationship between the annual and monthly standard
97 normal values. Instead, a unique regression model for each month was fit, which estimated that
98 months streamflow percentile based on a number of potential predictors. This approach was
99 flexible, as any variable that covers the reconstruction period could be used as a predictor.

Two APR models are considered in this study. The first uses only the standard normal of the lagged percentiles from the annual streamflow reconstruction:

$$\hat{N}_{m,0} = \beta_{-1}N_{a,-1} + \beta_0N_{a,0} + \beta_{+1}N_{a,+1} \quad (3)$$

where $N_{m,0}$ represents monthly standard normal, as in Eq. 2, and $N_{a,-1}$, $N_{a,0}$, and $N_{a,+1}$ are annual normalized percentiles for the previous, concurrent, and future water years, respectively. The associated β values are model coefficients fit using the regression approach described below. Transformation to standard normal values, rather than percentiles, was done to permit regression with standard, normal residuals. The model described in Eq. 3 is referred to as the APR model with lagged predictors. It can be viewed as a temporal disaggregation of the original annual reconstruction, using no additional data. The lagged regression approach alleviates the need to explicitly define the most appropriate water-year bounds, as it is naturally derived based on predictive power.

The second variant of the APR model incorporated predictors beyond the original annual reconstruction. This model, termed APR with all predictors for this study, was of the form:

$$\begin{aligned} \hat{N}_{m,0} = & \beta_{-1}N_{a,-1} + \beta_0N_{a,0} + \beta_{+1}N_{a,+1} + \beta_{ENSO}ENS O \\ & + \beta_{PDO}PDO + \beta_{PC1}PC1 + \dots \end{aligned} \quad (4)$$

where additional predictors for ENSO, PDO, and all regional tree-ring principal components are included as β_{ENSO} , β_{PDO} , and $\beta_{PC1} + \dots$, respectively.

Regression fitting was performed using elastic nets (Zou and Hastie, 2005), an alternative to Ordinary Least Squares (OLS), designed as an improvement over step-wise regression. Elastic nets are designed to avoid overfitting while handling high dimensional data with multicollinearity. This is accomplished by blending ridge regression (Hoerl and Kennard, 1970) and LASSO (least absolute shrinkage and selection operator) regression (Tibshirani, 1997, 1996) using a term, α , that can range from 0 (ridge regression) to 1 (LASSO regression). Blending these models combines the benefits of each: ridge regression is effective at reducing overfitting and handling highly related predictors, while the LASSO approach can perform variable selection by shrinking non-significant coefficients to zero. Ridge and LASSO regression use similar approaches that seek to minimize the sum of mean squared error along with a penalty term. The penalty terms are the primary difference between the approaches, using either the square of model

coefficients, $\sum \beta^2$, or the absolute value of the model coefficients, $\sum |\beta|$, for ridge and LASSO regression, respectively. The penalty is then multiplied by a shrinkage parameter, λ . For this study, optimal values of the blending term, α , and the shrinkage term, λ , were determined by repeated 10-fold cross-validation for each monthly model, in an effort to further avoid overfitting.

3. Methods

3.1. Mean Annual Reconstructions

The monthly reconstruction approaches were applied at two sites with existing annual streamflow reconstructions within the Bear River watershed in northern Utah. Flows in this region are snowpack-dominated, with the majority of annual precipitation falling in the form of winter snow. The two sites were located on the Logan River and in the upper Bear River. Both reconstructions (Allen et al., 2013; DeRose et al., 2015) used similar methods to reconstruct the MAF based on the USGS water-year of Oct-Sep.

The Bear River is located in the Intermountain U.S. and is the single largest river in the eastern Great Basin and drains portions of northeastern Utah, southwestern Wyoming, and South-eastern Idaho. The river headwaters are in the Uinta Mountains and the Bear River is the largest tributary to the Great Salt Lake, providing approximately 60% of its annual inflow. The annual flow reconstruction for the upper Bear River is based on the instrumental flow record from USGS gauge 1011500 (1942 CE-present). This gauge is located near the furthest upstream Utah-Wyoming border and includes flow contribution from a 445 km² watershed along the north slope of the Uinta mountains. Flow at this site contributes approximately 8% of the total Bear River flow. Flow at this site is considered near-natural, as it is upstream of major diversions, with only a single, small storage reservoir located upstream. The tree-ring predictor data used to reconstruct MAF was based on a nearby Utah juniper (*Juniperus osteosperma*) chronology. The reconstruction (DeRose et al., 2015) covered the period 800-2010 CE, and explained 67% of annual variation ($R^2 = 0.672$).

The Logan River is the largest tributary to the Bear River and drains 1389 km² of the Bear River Range in northeastern Utah and southeastern Idaho. The MAF reconstruction (Allen et al., 2013) is based on USGS gauge 10109001 (1922 CE - present), which captures the near-natural upper portion of the Logan River (554 km²) prior to impoundments and diversions near the city of Logan, Utah. The reconstruction covers the period from 1605 to 2010 CE. Two reconstruction

models were used, one termed the Local Model that relied on two Rocky Mountain juniper (*Juniperus scopulorum*) chronologies located within the Logan River watershed and another termed the Regional Model that also incorporated a two-needle pinyon (*Pinus flexilis*) chronology from north-central Utah and a limber pine (*Pinus flexilis*) chronology from western Wyoming. The Regional Model showed better skill ($R^2 = 0.581$) than the local model ($R^2 = 0.482$), but both reconstructions were considered for this study.

3.2. Global Circulation Index Predictors

Global circulation indices are one of the variables considered for inclusion in the APR model. The particular circulation indices included in this study were the El Niño-Southern Oscillation (ENSO, Trenberth, 1997) and the Pacific Decadal Oscillation (PDO, Mantua and Hare, 2002), which have been shown to affect the climate of North America, and Utah in particular (Cayan et al., 1999; Schoennagel et al., 2005; Wang et al., 2009b,a). Two ENSO reconstructions are available, based on different input data. The first ENSO reconstruction, described in detail by Li et al. (2011), is based on the leading Principal Component (PC) calculated from tree-ring chronologies developed for the North American Drought Atlas (Cook et al., 1999). This approach uses the documented effect of ENSO-related changes in atmospheric circulation on American climatology to reconstruct ENSO conditions in the Pacific Ocean back to 900 CE. In comparison, a more recent ENSO reconstruction (Li et al., 2013) is instead based on tree-ring chronologies from the tropics and mid-latitudes surrounding the Pacific Ocean (Asia, New Zealand, and North America, and South America). This provides a more comprehensive measure of conditions in the Pacific Ocean, where ENSO is defined, but covers a shorter period, only extending back to 1300 CE. Both reconstructions were considered as predictors, in an effort to determine which input provides more relevant data for the Bear River watershed. Consideration of both reconstructions was assumed to be reasonable due to their independent proxy data and a regression approach that can adequately handle highly correlated variables. Only one reconstruction of the PDO was considered for this study (Biondi et al., 2001), which relied on tree-ring chronologies along the west coast of the United States and Mexico, and extended to 1661 CE.

3.3. Tree-Ring Chronology Analysis

The final set of potential reconstruction predictor variables are based on regional tree-ring chronologies (49 chronologies located within a maximum 450 km radius of the reconstruction

gages). All tree-ring sites had limited soil development and occurred on southerly or westerly aspects, occupying open canopy positions. Chronologies included seven different species: Douglas-fir (*Pseudotsuga menziesii*), two-needle pinyon (*Pinus edulis*), singleleaf pinyon (*Pinus monophylla*), limber pine (*Pinus flexilis*), ponderosa pine (*Pinus ponderosa*), Utah juniper (*Juniperus osteosperma*), and Rocky Mountain juniper (*Juniperus scopulorum*). Tree-ring indices (ie, chronologies) were developed using the same approach for each site using the dplR package (Bunn, 2008). Individual tree-ring series were detrended using a cubic-smoothing spline with a frequency cut-off of 50% at a wavelength of $\frac{2}{3}$ of the length of the time-series to remove variability potentially associated with stand dynamics or biological growth trends. Tree-ring series were then averaged using a robust biweight mean. Autocorrelation was retained in the resultant tree-ring chronologies. Preliminary analysis (not presented) showed that correlation between monthly flow percentiles and different tree species chronologies peaked at different times of the year, suggesting that multiple species could be used to highlight unique portions of the seasonal hydroclimatic signal. Based on this, and the assumption that spatial climate patterns are important, particularly for large or heterogeneous watersheds, regional tree-ring chronologies were also considered as predictor variables.

[Figure 1 here.]

Regional tree-ring chronologies were selected to include a range of species, elevations, and locations in an effort to highlight the most important factors for predicting seasonal flow. Because tree-ring chronologies can be highly correlated, each responding to broad changes in water availability, Principal Component Analysis (PCA) was used to isolate unique differences in response, reducing the chronologies to a smaller set of orthogonal Principal Components (PCs) that explain a majority of the original tree-ring variance. Use of PCA for dimension reduction with highly correlated data is common in both the fields of dendrochronology and climatology. The number of retained PCs was based on the Kaiser-Guttman criteria (Kaiser, 1960; Guttman, 1954), defined as PCs with eigenvalues greater than 1, and by identifying a significant change in variance explained through a scree plot. These stopping criteria are thought to slightly overestimate the number of retained components (Jackson, 1993), which is desirable in the context of maximizing seasonal variability across multiple chronologies.

The regional tree-ring chronologies have varying lengths, primarily due to differences in their earliest measurements (800 - 1763 CE) (Fig. 1). Traditional PCA requires a fully defined data

matrix and therefore would require a trade-off, either removing short chronologies to lengthen the time series or producing PCs beginning in 1763 CE based on the common period. To avoid this limitation, while also capturing potentially valuable information in the shorter chronologies, missing values were imputed using the missMDA package (Josse, Julie and Husson, Franois, 2016) in R prior to PCA analysis. The number of components was determined by 10-fold cross-validation, minimizing mean squared error (MSE) across 100 simulations (Josse and Husson, 2012). Missing values were imputed using an iterative algorithm which cycles between estimating missing values and recalculating PCs until the observed and fitted values converge. PC scores were calculated in this way for a 600 year duration and compared to loadings derived from traditional PCA using approximately 400 (1610 CE, 36 chronologies) and 600 year (1390 CE, 19 chronologies) subsets , selected based on inflection points in chronology availability (Fig. 1). A third cut-off point at 900 years was initially considered, but dismissed because the number of available time series (4) was less than the optimal number of PCs to retain. The imputed PCA results did not significantly change results and was therefore used for all subsequent analysis. The Logan and Bear River sites were located close enough to each other that the same PCA scores were used for both locations.

3.4. Model Fitting and Goodness-of-Fit

Each of the model variants were fit using USGS stream gauge data, calculated by accumulating daily flows to determine mean monthly flow. The entire observed period (7/1942 to 9/2016 for Bear River and 10/1921 to 9/2015 for Logan River) was used as the dependent variable for fitting the models. For each site, two reconstruction models were generated, the first using the true observed MAF as training data and the second using the reconstructed MAF. The model using observed MAF provides an upper envelope for model performance, given perfect input data, while the model using reconstructed MAF provides an estimate of model fit that can be expected for the full reconstruction.

Goodness-of-fit was evaluated using a suite of tests and measures during the common reference period of 10/1942 to 9/2015, to allow for equal comparisons across the two sites. These goodness-of-fit measures included mean error (ME) to estimate systematic model bias, mean absolute error (MAE) and root-mean-squared -error (RMSE) to estimate the residual magnitudes, measures of parametric (Pearson, R) and non-parametric (Spearman, $R_{Spearman}$) correlation, and Nash-Sutcliffe Efficiency (NSE) to estimate overall model predictive power. In addition, residual

244 and time-series plots were generated to allow for visual inspection of residual patterns or timing
245 issues that might not be captured by these metrics (McCuen et al., 2006; Criss and Winston,
246 2008). Finally, model fit was evaluated for each month separately using the same goodness-of-fit
247 metrics and residual plots. This allowed for a more detailed review of model performance, par-
248 ticularly in months with low flow, which could otherwise be obscured by high flow periods in the
249 full time series.

250 **4. Results**

251 *4.1. Monthly Fraction (MF) Model Fitting*

252 Estimation of the MF model required only the calculation of the mean monthly proportion of
253 flow. The Logan and Bear River sites showed similar seasonal patterns, with strong seasonality
254 and flow peaks in the summer caused by delayed melting of the winter snowpack (Fig. 2). Figure
255 2 presents all available years, with the mean monthly fraction used for reconstruction shown as
256 a dark line. Logan River (Fig. 2a) had a higher proportion as baseflow and more gradual rising
257 and falling limbs around a peak which occurred in May or June. The Bear River (Fig. 2b) had a
258 more drastic seasonal peak flow, that predominantly occurred in June and occasionally May.

259 [Figure 2 here.]

260 *4.2. Annual Percentile (AP) Model Fitting*

261 The primary model fitting task for the AP model was to determine the best cumulative prob-
262 ability distribution for the MAF and each of the monthly flows. For both sites, skewness-kurtosis
263 plots recommended the use of a gamma distribution for monthly flows and a logistic distribu-
264 tion for MAF. Choice of these distributions was further supported by low AIC values and non-
265 significant tests for violations of the candidate empirical cumulative distributions. The logistic
266 distribution used for annual flows was nearly Gaussian, but with slightly thicker tails, while
267 the gamma distribution used for monthly flows accounted for the positively skewed nature of
268 monthly flows combined with a lower bound at zero.

269 *4.3. APR Model with Lagged Predictors*

270 The APR model with lagged annual flows reconstruction was equivalent to a temporal dis-
271 aggregation of the original reconstruction and required model coefficients for the effect of the

lagged (-1, 0, +1) annual standard normal values on each of the monthly normalized percentiles. The resulting fit (Fig. 3) showed that the importance of the reconstructed annual flow, quantified by the coefficient, transitions between the reconstructed flow from the previous year (-1) to the concurrent year (0) for both locations. For the Logan River (Fig. 3a), this transition was gradual, where the crossover point occurred between February and March, and the highest values occurred during the peak and falling limb of the seasonal hydrograph (May through September), before being carried over into the next water-year. The Bear River had a much more abrupt transition (Fig. 3b) between the previous water-year, which explained October through March, and the concurrent water year, which explained monthly flow from June through the end of the water year. As expected, the future (+1) water-year reconstruction had little to no predictive power for monthly flows.

[Figure 3 here.]

4.4. Tree-Ring Chronology Analysis

From the imputed regional tree-ring dataset, which included all 49 chronologies (Fig. 1) and thereby the widest range of site conditions, 8 PCs were retained based on minimizing MSE from repeated 10-fold cross-validation. This result was similar to the result from a traditional PCA and a smaller (36 chronologies) dataset, which recommended retaining 8 PCs based on the Kaiser-Guttman test. As expected, the smaller 600-year, 19-chronology dataset had a smaller range of species and sites, which resulted in the retention of 5 PCs. Retaining 8 PCs with imputation was therefore deemed reasonable.

In all tests, the first PC (PC1) explained a large proportion of the total variance, from 38.4% in the imputed dataset to 33.9% and 35.7% in the 400- and 600-year datasets, respectively. The proportion of variance explained by subsequent PCs fell after PC1, from 7.8% to 2.7% between PCs 2 and 8, and ultimately explained 73.1% of the total variance. Loadings for the imputed dataset were nearly identical to loadings from the traditional PCA based on spatial and species patterns, which further validated the imputation scheme.

While a total of eight PCs were considered as predictors, focus was given to the first five, as these represented a majority of the explained variance. PC1 was the only PC where all chronologies loaded in the same direction (Fig. 4). Because of this common pattern and the relatively large variance explained, PC1 appeared to characterize the general dry or wet state of the en-

302 tire region. Loading for PC1 was dominated by Douglas-fir from low-elevation sites (Fig. 4).
303 However, this does not mean that Douglas-fir was necessarily best at capturing the regions hydro-
304 climatology. For example, both annual reconstructions (Allen et al., 2013; DeRose et al., 2015)
305 originally considered Douglas-fir chronologies, but instead relied on other species that better ex-
306 plained MAF. Instead, Douglas-fir likely dominated PC1 because they have good predictive skill
307 and they were the best represented species (43%) among the available chronologies.

308 [Figure 4 here.]

309 While PC1 explained regional hydroclimatology and a major portion of the variability among
310 all chronologies, subsequent PCs captured smaller spatial and species-specific patterns and devi-
311 ations from the larger climate signal (Fig. 4). PC loadings are shown spatially and by species in
312 Fig. 4 and were summarized as:

- 313 • PC1: General hydroclimatology, as captured by Douglas-fir at lower elevations
- 314 • PC2: Dipole between Bear River Douglas-fir and Southern Utah pinyon/ponderosa pine
- 315 • PC3: High-elevation sites, dominated by two limber pine and two Douglas-fir chronologies
- 316 • PC4: East/west dipole between Utah juniper species and Wyoming Douglas-fir/limber
317 pines
- 318 • PC5: Dipole between juniper sites near the Logan River and a singleleaf pinyon site in the
319 Great Basin.

320 Subsequent PCs (shown in the appendix) become more site-specific and explained less vari-
321 ance.

322 4.5. APR Model with All Predictors

323 The most flexible and complex model considered in this study was the APR model using all
324 predictors, including lagged annual flow reconstruction, global circulation indices, and regional
325 tree-ring PCs. Model coefficients for lagged annual flow reconstruction had nearly identical
326 shapes as those presented in Section 4.3 and Figure 3, with a transition from the prior years
327 reconstruction to the current year, crossing in spring (March-May). This transition was more
328 gradual for the Logan River and more dramatic for Bear River.

329

[Figure 5 here.]

320 The two sites differed in their response to global circulation indices. Monthly flow for the
 321 Logan River could be partially predicted by ENSO, where positive ENSO values indicated de-
 322 creased flow during the winter and spring (January to May) and vice-versa (Fig. 5a). The best
 323 predictor shifted from ENSO, as defined by Pacific Ocean proxies Li et al. (2013) during the
 324 winter, to ENSO as quantified by tree-rings from the North American Drought Atlas (NADA)
 325 (Li et al., 2011). For the Bear River, PDO, rather than ENSO, was a significant predictor of flow.
 326 This predictive skill occurred during the late summer and fall (August through December), peak-
 327 ing in September (Fig. 5c). The sign of the coefficient suggest that late summer and fall flows in
 328 the Bear River are higher than typical when the PDO is strongly positive, and vice-versa.

330 At both sites, PC1 had an insignificant effect when considered as a predictor within the full
 331 APR model. The most influential PC for both sites was PC2 during the summer (July-September)
 332 and PC6 during the winter (December-April) (Fig. 5b-d). Results for PC2 suggest that summer
 333 flows are slightly higher than would be otherwise predicted if southern pines signal wetter con-
 334 ditions than northern Douglas-firs. The PC6 response shows that large tree-ring growth at two
 335 high elevation (2000, 2350 m) sites along the Wyoming border leads to greater winter flows. The
 336 Logan River has a unique effect, with PC8 providing some predictive power in the Logan River
 337 during winter. The remaining PCs have significantly less predictive power or vary considerably
 338 from month to month, suggesting their contribution is statistical noise rather than a true driver
 339 of the flow signal. Lagged PC effects were initially considered, using lags of -2 to +2 years, but
 340 were ultimately removed due to weak explanatory skill and concern about over-fitting.

350 *4.6. Goodness-of-Fit*

351 The APR model with all predictors produced the best fit across all goodness-of-fit measures,
 352 which included analyses of the full time-series, monthly fit, residual patterns, and visual in-
 353 spection. Considering the full time-series, goodness-of-fit improved across nearly all measures
 354 (MAE, RMSE, NSE, R) with increasing numbers of parameters, except for mean error (Table 1).
 355 Correlations (R) for the APR model with all predictors were 0.97 for both sites. This represents
 356 a theoretical best-fit, assuming observed annual flows were known. In all cases, using annual
 357 flows reconstructed from tree-rings lowers the goodness-of-fit, decreasing R to 0.94 and 0.88,
 358 again for the Bear and Logan Rivers. For many fit measures, as additional parameters are added

359 to the model, not only does the overall fit improve, but the gap between observed annual flows
360 and reconstructed annual flows decreases. The MF model is unique in that it excels at minimizing
361 ME, i.e. systematic bias, but it tends to perform only moderately well to poorly for other measures.

362 Goodness-of-fit at the monthly scale provided more detailed information about model perfor-
363 mance. Using the NSE as a goodness of fit measure, it is clear that the MF and AP models have
364 no predictive power (NSE 0) for much of the early water year, from October until late winter or
365 spring (Fig. 6). The MF model performs particularly poorly during this period. There is a rapid
366 increase in model skill during May and June, when the majority of annual flow volume occurs
367 (Fig. 2). The MF and AP models perform similarly during this period, with the MF model per-
368 forming slightly better for reconstructed annual flows (Fig. 6b,d). NSE remained high for the
369 remainder of the water year.

370 [Figure 6 here.]

371 [Table 1 here.]

372 Removing the pre-defined water year in the APR model with lagged annual reconstruction
373 predictors represents the largest structural improvement among the models (Fig. 7). The im-
374 provement by the APR model occurred primarily during the early part of the water year, allowing
375 for near-constant NSE throughout the year 7). The only exception to this rule occurs in April,
376 where a characteristic poor fit occurs across all models.

377 [Figure 7 here.]

378 Expanding the APR model to include global climate indices and regional tree-rings produced
379 smaller, but still significant improvements in model fit. The largest improvements occurred in the
380 Bear River, particularly during the rising (May) and recession (July through September) limbs of
381 the seasonal hydrograph (Fig. 6d). If the model predictors are added successively (Fig. 7), global
382 climate indices produce the larger improvement, while regional tree-rings provide the smallest
383 incremental improvement.

384 A sample of the historical reconstruction for Logan River is presented in Fig. 8a, highlighting
385 a period of severe regional drought during the 1730s (Woodhouse and Brown, 2001). The en-
386 tire reconstructed timeseries is available online at https://jstagge.shinyapps.io/paleo_

387 flow. By instead focusing on the observed period and visually comparing reconstructed to ob-
388 served flows (Fig. 8b), it was possible to confirm the good fit of the APR model and to identify
389 structural errors in the MF and AP models. In these models, the transition between water years
390 was not smooth and sometimes produced drastic jumps in flow, particularly between years with
391 very different annual reconstructions, such as between 1961 and 1962 (Fig. 8b). This shift from
392 a dry to wet year produced a sudden increase in flow at the water year transition, prior to the high
393 flow period recorded in the observations. This type of error was not present in the APR models.
394 In addition to smooth annual transitions, including all predictors in the APR model improved the
395 fit and shape of the rising and falling hydrograph limbs (Fig. 8b), despite only producing a mod-
396 est improvement in monthly and annual fit metrics. This finding is further supported by reduced
397 residuals for medium and low flows in the non-peak season when using regional tree-rings and
398 global climate indices. All examined models tend to underestimate maximum flows, particularly
399 when using reconstructed MAF, because the annual reconstruction also tends to underestimate
400 high flows.

401 [Figure 8 here.]

402 **5. Discussion**

403 *5.1. Reconstruction Model Comparison*

404 Of the models proposed in this study, the APR model provides the most favorable mea-
405 sures of goodness-of-fit. This model produces NSE greater than zero (0.26-0.60) for all months
406 except April and Pearson's correlation coefficients (R) are 0.94 and 0.88 for the Bear and Lo-
407 gan Rivers, respectively, confirming that the model can adequately reproduce monthly reference
408 period flows. Goodness of fit decreases slightly when using reconstructed MAF, rather than ob-
409 served MAF, but this decrease is minor and reasonable given the challenge and potential benefit
410 of reconstructing centuries of streamflow at a monthly scale.

411 The MF model is the simplest approach, but it makes a strong assumption about seasonal
412 flows, namely that the hydrograph shape is constant and thus independent of annual flow. Be-
413 cause of this assumption, the MF model accurately captures flows during the peak, but tends to
414 produce errors in timing and magnitude for all other parts of the year. This is particularly an issue
415 for low flows, which are scaled linearly with annual flow. This assumption is not reasonable be-
416 cause groundwater sources provide a buffer to low flows. Additionally, the MF model produces

417 artifacts at the transitions between water years. The only metric for which the MF model excels
418 is mean error, a measure of consistent model bias. This is because the MF model relies on the
419 seasonal mean proportion, which in turn ensures that reconstructed flows are evenly distributed
420 around mean flows, even if accuracy is poor.

421 The AP model addresses the issue of hydrograph shape, allowing the shape to change with
422 annual flow percentile, but it retains the limitation of applying the same percentile across the en-
423 tire water-year. This also produces calculation artifacts at the water-year transitions and scaling
424 errors that propagate backwards in time from the peak season. This results in little or no predic-
425 tive power prior to the peak flow season. Another potential limitation of the AP approach is the
426 potential for generating negative flows if the univariate distribution used to transform flows does
427 not have a lower bound. Here, we have prevented this by selecting a distribution that is strictly
428 positive.

429 The introduction of regression in the APR model provides the most significant improvement
430 in reconstruction quality. With this improvement, the gap in NSE decreased between models
431 using reconstructed MAF and observed MAF, considered the feasible upper limit for fit. For
432 the sites examined here, this allowed fall and winter low flows to retain information from the
433 prior flow peak, rather than depending on the peak flow that was yet to occur. Adding additional
434 predictors, such as global climate indices and regional tree-ring signals provides further model
435 improvement, which although minor remained statistically significant. Global climate indices
436 were the second most important set of predictors, followed by regional tree-rings. Including
437 these additional predictors may be a trade-off because data availability may limit the feasible
438 length of reconstructions. The approach provided here attempted to minimize this limitation
439 through a process of PCA imputation.

440 It should be noted that correlation (R) values shown here for the full time series are not
441 directly comparable to R values from annual reconstructions because some of the variance ex-
442 plained is due to capturing the seasonal pattern of flows. Regardless, the R values are exception-
443 ally good and further supported by excellent model skill measured for individual months.

444 Differences in goodness of fit between the two sites can partially be attributed to differences
445 in their hydrologic response. Reconstructions at the Logan River had greater explanatory power
446 throughout the year, which might be due to its larger area and higher baseflow, which allow for
447 longer hydrologic “memory”. The Bear River, with its smaller area and greater dependence on

448 snowmelt produced a flashier signal, which made reconstruction more difficult. Reconstructions
449 in this flashy watershed were best for the peak flow period (June through September).

450 *5.2. Interpretation of Model Predictors*

451 In addition to providing the best model fit, the inclusion of external predictors in the APR
452 model provides an opportunity to explore the effect of global teleconnections and regional hy-
453 droclimatology on local streamflow as a form of validation. However, it is important to note this
454 study was not designed to test these links or to isolate their effects. All predictors were consid-
455 ered simultaneously in the APR model and therefore, if the effect of a given driver is accounted
456 for by another measure, it would not appear to be significant. For instance, if the effect of ENSO
457 is already accounted for in the annual MAF reconstruction or modeled better by the regional
458 patterns of tree-rings, it would not appear in the final model.

459 In the western U.S., previous studies have shown that a positive ENSO index is associated
460 with a warmer, drier winter in northern states and a wetter, cooler winter in the south (Cayan
461 et al., 1999; Redmond and Koch, 1991). Our sites are situated near the fulcrum of these opposed
462 climatic drivers. The APR model results were consistent with these studies (Cayan et al., 1999;
463 Redmond and Koch, 1991; Kurtzman and Scanlon, 2007), where the Logan River gage, located
464 farther north than the Bear River gage, exhibited a relationship to ENSO, while the Bear River did
465 not. For the Logan River gage, positive ENSO decreased winter flows, consistent with predicted
466 ENSO effects on northern regions, and with model improvements found by Allen et al. (2013).
467 The best predictor of this ENSO effect transitioned between the two available measures of ENSO
468 strength, with trees throughout the Pacific Ocean region providing more predictive skill during
469 winter and North Atlantic trees providing more skill during spring. This follows logically from
470 the difference in growing season between the northern and southern hemispheres.

471 The lack of a significant ENSO effect in the Bear River was curious, as it is located relatively
472 close to the Logan River site. The Bear River headwater gage has a unique orientation along
473 the north slope of the Uinta Mountains, the most prominent east-west-oriented mountain in the
474 United States, and which exceeds elevations of 4,000 m. This large mountain range receives
475 substantial winter snow due to orographic effects on Pacific westerlies. It also receives summer
476 precipitation, mostly due to convective storms, because storm tracks originating from the south
477 can be blocked by or redirected east by the prominent range. That the Bear River site was
478 influenced more by the PDO during the summer than ENSO, is likely due to PDO's ability

479 to explain variability unaccounted for in ENSO. While a previous analysis of streamflow on
480 the Ashley River (MacDonald and Tingstad, 2007), also located in the Uinta Mountain range,
481 showed a weak but significant correlation between the PDO and discharge (0.24), corroborating
482 our findings, the cause of this relationship requires future scrutiny.

483 The regional tree-ring predictors were useful for validating model results. PC1 was likely
484 not important for explaining regional wet/dry variability because this measure of overall wet-
485 ness is already captured by the annual flow reconstructions in the model. PC2 was an important
486 predictor for monthly summer flow at both sites. The spatial pattern of PC2 mirrors the ENSO
487 dipole, where northerly sites (predominantly Douglas-fir) responded in an opposite manner to
488 more southerly species, which were primarily two-needle pinyon, but also ponderosa pine and
489 Douglas-fir. This counterintuitive result is best explained by considering that the annual stream-
490 flow reconstruction already explains much of the variability associated with tree-ring sites located
491 near the gages, leaving the north-south chronology dichotomy to mediate baseflow recession, and
492 prevent flows from decreasingly excessively during dry years. PC3 directly captured variability
493 in ring width at high-elevation sites and should therefore reflect variation in monthly streamflow
494 not predicted by the more prevalent low-elevation sites. PC4, though not a dominant pattern
495 in the reconstruction models, indicated an east/west dipole identified in previous studies (Wang
496 et al., 2009b).

497 The APR model had difficulty modeling flows at both sites during the months of April and
498 May. During these months, all coefficients for external predictors were either set to zero or the
499 coefficients became large and opposed. These are the two opposed behaviors of LASSO and
500 ridge regression, respectively, within the elastic net framework as one or the other attempts to
501 handle poor predictive skill. For the Bear River in particular, none of the time series showed any
502 predictive capabilities in April, including the lagged annual reconstruction. This is likely because
503 April and May occur at the seasonal beginning of the snowmelt season when minor changes in
504 temperature near 0°C and solar radiation can produce large changes in snowmelt, runoff, and
505 flow. Rapid changes in temperature that likely drive snowmelt timing and amount are unlikely
506 to be directly captured in growing season ring width indices such as those used in this study. For
507 future implementations of the APR model, it will be important to consider additional climatic
508 predictors to help improve the explanation of spring snowmelt timing.

509 *5.3. Implications for Water Management*

510 The model presented here is the first approach to reconstruct monthly streamflows directly
511 from tree-ring chronologies and climate reconstructions, without routing climate reconstructions
512 through a hydrologic model. With reasonable estimates of monthly flow, extending many cen-
513 turies back in time, water managers can challenge systems models with a larger range of natural
514 variability in the duration and severity of drought and pluvial events. This is a major step for-
515 ward for the use of paleoclimate data in water management applications, permitting better water
516 supply forecasts, simulation, and more robust optimization of reservoir operations. For example,
517 engineers are often tasked with the seemingly incongruous and uncertain task of designing sys-
518 tems to address drought events with 100-year return periods in watersheds where the observed
519 flow record is shorter than 100 years. Extrapolation using extreme value analysis can estimate the
520 magnitude of a hypothetical event, but multi-century flow reconstructions include more useful
521 and realistic scenarios. Multi-year or decadal periods of drought and successive small drought
522 events can challenge water management operations and reservoir recovery in ways that a single,
523 severe drought event does not. Establishing a long catalog of near-natural flow is also critical
524 when estimating future hydrologic risks under conditions of a non-stationary climate.

525 One other methodology has been developed to address the need for monthly reconstructed
526 streamflows (Gangopadhyay et al., 2015), although it does not produce flow reconstructions di-
527 rectly from tree-ring chronologies. Instead, this method uses resampling (Gangopadhyay et al.,
528 2009) to generate many equally likely time series of monthly temperature and precipitation,
529 which are then used as inputs to a water balance model. This approach therefore has the benefit
530 of ensuring water balance within large watersheds and producing many ensembles to be used for
531 water systems model simulations. However, this resampling approach has several limitations not
532 present in the APR model. For example, by resampling full years, the number of potential tem-
533 perature and precipitation time series are limited to the years in the observed record. For example,
534 if applied to the Logan and Bear Rivers, there would be 88 and 68 potential annual segments,
535 respectively. Using k=10 nearest neighbor resampling, the same years would be repeated across
536 many realizations. This ensures realistic years, but severely limits variation within the range of
537 observed conditions and prevents reasonable extrapolation slightly outside this range. Addition-
538 ally, by requiring a fully-developed water balance model, this approach limits its applicability
539 for regions with adequate gauge data and modeling.

540 The APR model proposed here has both greater flexibility and lower data requirements,
541 avoiding the need for developing a full hydrologic model, while also allowing a freer range
542 of flows, still constrained by the historical flow distribution. The APR model also provides a
543 direct statistical link between tree-ring proxies and global circulation drivers that can be used for
544 model validation. While it is our contention that the model outlined here provides several impor-
545 tant improvements over the approach of Gangopadhyay et al. (2015), it is also important to note
546 that these two methods are fundamentally different and are designed to address different research
547 questions. Because of this difference in concept, the Gangopadhyay et al. (2015) method focuses
548 on capturing uncertainty and ranges of potential flow, while the method proposed here focuses
549 on a single best estimate, making direct goodness of fit comparisons difficult. Given the value
550 of monthly streamflow reconstructions for water managers and the potential growth in this new
551 field of research, it is important for future studies to test the global applicability of both methods
552 by evaluating their use outside the western U.S.

553 **6. Conclusions**

554 A new framework for generating monthly streamflows directly from annual reconstructions
555 was introduced and demonstrated using two sites in northern Utah. Four potential models were
556 evaluated, beginning with a simple monthly flow ratio (MF model), introducing a link between
557 annual and monthly percentiles (AP model), allowing lagged annual flows to predict monthly per-
558 centiles (APR model), and ultimately including regional and global indices as regressors (APR
559 model with predictors). Of the potential variants of this framework, the APR model was shown to
560 successfully reproduce flows across nearly all months and across the entire range of flows. This
561 model can be applied using only data from annual flow reconstructions or can incorporate re-
562 gional tree-ring chronologies and global climate index reconstructions. Adding these additional
563 predictors was shown to improve predictive skill, while also providing insight with regard to
564 streamflow drivers. More simple model variants were presented as a comparison and to highlight
565 potential modeling challenges, such as disconnections at water year transitions.

566 Paleostreamflow reconstructions have not been readily incorporated into water systems plan-
567 ning because no methods existed to accurately reconstruct monthly flows until recently. Monthly
568 flows, rather than annual flows, are critical for water decisions and planning. The method
569 described in this study makes monthly paleostreamflow reconstruction more tractable, putting

570 these products in the hands of water managers, systems analysts, and decision makers. To date,
571 monthly reconstructions have only been generated using case studies in the western U.S. More
572 study is needed to determine whether these approaches can be applied globally and whether there
573 are additional climate proxies that could be used to improve reconstructions.

574 **Acknowledgements**

575 This work was funded by Utah Mineral Lease funds.

576 This article was prepared in part by an employee of the USDA Forest Service as part of
577 official duties and is therefore in the public domain.

578 We would like to acknowledge the many dendrochronologists who provided tree-ring data to
579 the International Tree-Ring Data Bank: Stephen Gray, Erika Wise, Connie Woodhouse, and Jeff
580 Lukas.

581 **References**

- 582 Allen, E.B., Rittenour, T.M., DeRose, R.J., Bekker, M.F., Kjelgren, R., Buckley, B.M., 2013. A tree-ring based re-
583 construction of Logan River streamflow, northern Utah. *Water Resources Research* 49, 8579–8588. doi:10.1002/
584 2013WR014273.
- 585 Biondi, F., Gershunov, A., Cayan, D.R., 2001. North Pacific Decadal Climate Variability since 1661. *Journal of Climate*
586 14, 5–10. doi:10.1175/1520-0442(2001)014<0005:NPDCVS>2.0.CO;2.
- 587 Bonin, D.V., Burn, D.H., 2005. Use of tree ring reconstructed streamflows to assess drought. *Canadian Journal of Civil*
588 *Engineering* 32, 1114–1123. doi:10.1139/105-069.
- 589 Bunn, A.G., 2008. A dendrochronology program library in R (dplR). *Dendrochronologia* 26, 115–124. doi:10.1016/
590 j.dendro.2008.01.002.
- 591 Cayan, D.R., Redmond, K.T., Riddle, L.G., 1999. ENSO and Hydrologic Extremes in the Western United States. *Journal*
592 *of Climate* 12, 2881–2893. doi:10.1175/1520-0442(1999)012<2881:EAHEIT>2.0.CO;2.
- 593 Cook, E.R., Meko, D.M., Stahle, D.W., Cleaveland, M.K., 1999. Drought Reconstructions for the Continental United
594 States. *Journal of Climate* 12, 1145–1162. doi:10.1175/1520-0442(1999)012<1145:DRFTCU>2.0.CO;2.
- 595 Criss, R.E., Winston, W.E., 2008. Do Nash values have value? Discussion and alternate proposals. *Hydrological*
596 *Processes* 22, 2723–2725. doi:10.1002/hyp.7072.
- 597 Crutcher, H.L., 1975. A Note on the Possible Misuse of the Kolmogorov-Smirnov Test. *Journal of Applied Meteorology*
598 14, 1600–1603. doi:10.1175/1520-0450(1975)014<1600:ANOTPM>2.0.CO;2.
- 599 Cullen, A.C., Frey, H.C., 1999. Probabilistic techniques in exposure assessment: a handbook for dealing with variability
600 and uncertainty in models and inputs. Springer Science & Business Media.
- 601 Delignette-Muller, M.L., Dutang, C., 2015. fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical*
602 *Software* 64, 1–34. URL: <http://www.jstatsoft.org/v64/i04/>.

- 603 DeRose, R.J., Bekker, M.F., Wang, S.Y., Buckley, B.M., Kjelgren, R.K., Bardsley, T., Rittenour, T.M., Allen, E.B., 2015.
 604 A millennium-length reconstruction of Bear River stream flow, Utah. *Journal of Hydrology* 529, Part 2, 524–534.
 605 doi:10.1016/j.jhydrol.2015.01.014.
- 606 Gangopadhyay, S., Harding, B.L., Rajagopalan, B., Lukas, J.J., Fulp, T.J., 2009. A nonparametric approach for
 607 paleohydrologic reconstruction of annual streamflow ensembles. *Water Resources Research* 45, doi:10.1029/
 608 2008WR007201.
- 609 Gangopadhyay, S., McCabe, G.J., Woodhouse, C.A., 2015. Beyond annual streamflow reconstructions for the Upper
 610 Colorado River Basin: A paleo-water-balance approach. *Water Resources Research* 51, 9763–9774. doi:10.1002/
 611 2015WR017283.
- 612 Guttman, L., 1954. Some necessary conditions for common-factor analysis. *Psychometrika* 19, 149–161. doi:10.1007/
 613 BF02289162.
- 614 Hoerl, A.E., Kennard, R.W., 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*
 615 12, 55–67. doi:10.1080/00401706.1970.10488634.
- 616 Jackson, D.A., 1993. Stopping Rules in Principal Components Analysis: A Comparison of Heuristical and Statistical
 617 Approaches. *Ecology* 74, 2204–2214. doi:10.2307/1939574.
- 618 Josse, J., Husson, F., 2012. Selecting the number of components in principal component analysis using cross-validation
 619 approximations. *Computational Statistics & Data Analysis* 56, 1869–1879. doi:10.1016/j.csda.2011.11.012.
- 620 Josse, Julie, Husson, Franois, 2016. missMDA: A Package for Handling Missing Values in Multivariate Data Analysis |
 621 Josse]. *Journal of Statistical Software* 70. doi:10.18637/jss.v070.i01.
- 622 Kaiser, H.F., 1960. The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Mea-
 623 surement* 20, 141–151. doi:10.1177/001316446002000116.
- 624 Kurtzman, D., Scanlon, B.R., 2007. El Niño-Southern Oscillation and Pacific Decadal Oscillation impacts on precipi-
 625 tation in the southern and central United States: Evaluation of spatial distribution and predictions. *Water Resources
 626 Research* 43, W10427. doi:10.1029/2007WR005863.
- 627 Li, J., Xie, S.P., Cook, E.R., Morales, M.S., Christie, D.A., Johnson, N.C., Chen, F., D'Arrigo, R., Fowler, A.M., Gou,
 628 X., Fang, K., 2013. El Niño modulations over the past seven centuries. *Nature Climate Change* 3, 822–826. doi:10.
 629 1038/nclimate1936.
- 630 Li, J.B., Xie, S.P., Cook, E.R., Huang, G., D'Arrigo, R., Liu, F., Ma, J., Zheng, X.T., 2011. Interdecadal modulation of
 631 El Niño amplitude during the past millennium. *Nature Climate Change* 1, 114–118. doi:10.1038/Nclimate1086.
- 632 Loaiciga, H.A., Haston, L., Michaelsen, J., 1993. Dendrohydrology and long-term hydrologic phenomena. *Reviews of
 633 Geophysics* 31, 151–171. doi:10.1029/93RG00056.
- 634 MacDonald, G.M., Tingstad, A.H., 2007. Recent and Multicentennial Precipitation Variability and Drought Occur-
 635 rence in the Uinta Mountains Region, Utah. *Arctic, Antarctic, and Alpine Research* 39, 549–555. doi:10.1657/
 636 1523-0430(06-070) [MACDONALD]2.0.CO;2.
- 637 Mantua, N.J., Hare, S.R., 2002. The Pacific Decadal Oscillation. *Journal of Oceanography* 58, 35–44. doi:10.1023/A:
 638 1015820616384.
- 639 McCuen, R.H., Knight, Z., Cutter, A.G., 2006. Evaluation of the Nash-Sutcliffe Efficiency Index. *Journal of Hydrologic
 640 Engineering* 11, 597–602. doi:10.1061/(ASCE)1084-0699(2006)11:6(597).
- 641 Meko, D.M., Therrell, M.D., Baisan, C.H., Hughes, M.K., 2001. Sacramento River Flow Reconstructed to A.D. 869

- 642 from Tree Rings. JAWRA Journal of the American Water Resources Association 37, 1029–1039. doi:10.1111/j.
643 1752–1688.2001.tb05530.x.
- 644 Redmond, K.T., Koch, R.W., 1991. Surface Climate and Streamflow Variability in the Western United States and
645 Their Relationship to Large-Scale Circulation Indices. Water Resources Research 27, 2381–2399. doi:10.1029/
646 91WR00690.
- 647 Schoennagel, T., Veblen, T.T., Romme, W.H., Sibold, J.S., Cook, E.R., 2005. ENSO and PDO Variability Affect Drought-
648 Induced Fire Occurrence in Rocky Mountain Subalpine Forests. Ecological Applications 15, 2000–2014. doi:10.
649 1890/04-1579.
- 650 Solander, K., Saito, L., Biondi, F., 2010. Streamflow simulation using a water-balance model with annually-resolved
651 inputs. Journal of Hydrology 387, 46–53. doi:10.1016/j.jhydrol.2010.03.028.
- 652 Tibshirani, R., 1996. Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series
653 B (Methodological) 58, 267–288. URL: <http://www.jstor.org/stable/2346178>.
- 654 Tibshirani, R., 1997. The Lasso Method for Variable Selection in the Cox Model. Statistics in Medicine 16, 385–395.
655 doi:10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3.
- 656 Trenberth, K.E., 1997. The Definition of El Niño. Bulletin of the American Meteorological Society 78, 2771–2777.
657 doi:10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2.
- 658 Wang, S.Y., Gillies, R.R., Jin, J., Hipps, L.E., 2009a. Coherence between the Great Salt Lake Level and the Pacific
659 Quasi-Decadal Oscillation. Journal of Climate 23, 2161–2177. doi:10.1175/2009JCLI2979.1.
- 660 Wang, S.Y., Gillies, R.R., Jin, J., Hipps, L.E., 2009b. Recent rainfall cycle in the Intermountain Region as a quadrature
661 amplitude modulation from the Pacific decadal oscillation. Geophysical Research Letters 36, L02705. doi:10.1029/
662 2008GL036329.
- 663 Wolock, D., McCabe, G., 1999. Effects of potential climatic change on annual runoff in the conterminous United States.
664 J. Am. Water Resour. Assoc 35, e1350.
- 665 Woodhouse, C.A., Brown, P.M., 2001. Tree-Ring Evidence for Great Plains Drought. Tree-Ring Research 57, 89–103.
- 666 Woodhouse, C.A., Gray, S.T., Meko, D.M., 2006. Updated streamflow reconstructions for the Upper Colorado River
667 Basin. Water Resources Research 42, doi:10.1029/2005WR004455.
- 668 Woodhouse, C.A., Lukas, J.J., 2006. Multi-Century Tree-Ring Reconstructions of Colorado Streamflow for Water Re-
669 source Planning. Climatic Change 78, 293–315. doi:10.1007/s10584-006-9055-0.
- 670 Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical
671 Society: Series B (Statistical Methodology) 67, 301–320. doi:10.1111/j.1467-9868.2005.00503.x.

672 **List of Figures**

673 1	Available tree-ring chronologies.	26
674 2	Monthly flow proportion relative to mean annual flow (%) for Logan (a) and Bear 675 River headwaters (b).	27
676 3	APR model coefficients for concurrent and lagged (-1, +1) years at the (a) Logan 677 and the (b) Bear rivers.	28
678 4	PC loadings from imputed values presented spatially (left) and by species (right). 679 Loadings are shown for PCs 1-5 (a-e). Spatial maps only present sites with 680 loadings greater than the mean to prevent over-plotting.	29
681 5	Model coefficients for the Logan River (a-b) and the Bear River (c-d). Global 682 circulation indices are shown on the left (a,c), while tree-ring PCs are shown on 683 the right (b,d).	30
684 6	Nash-Sutcliffe Efficiency (NSE) by month.	31
685 7	Nash-Sutcliffe Efficiency (NSE) for the Logan River. Models of increasing com- 686 plexity are presented along the horizontal axis, while colors represent the input 687 reconstruction dataset.	32
688 8	Reconstructed flows at the Logan River site for subsets of the (a) historical and 689 (b) observed periods.	33

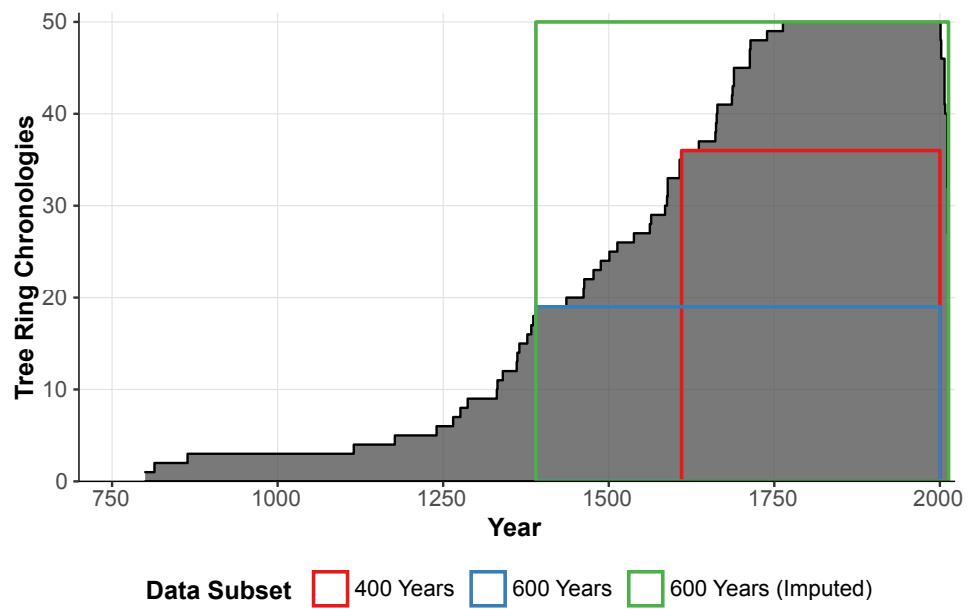


Figure 1: Available tree-ring chronologies.

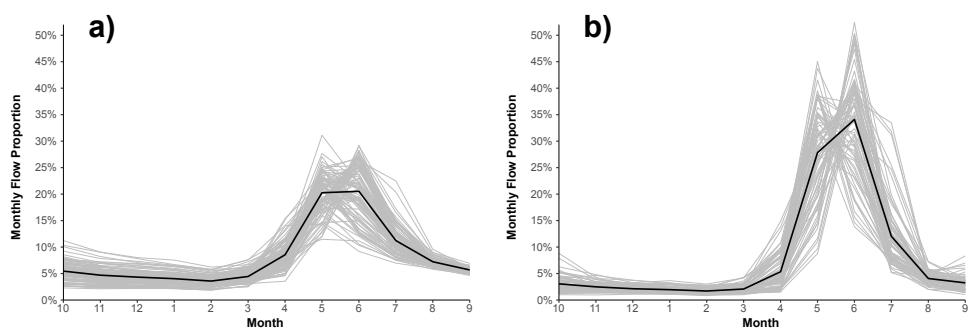


Figure 2: Monthly flow proportion relative to mean annual flow (%) for Logan (a) and Bear River headwaters (b).

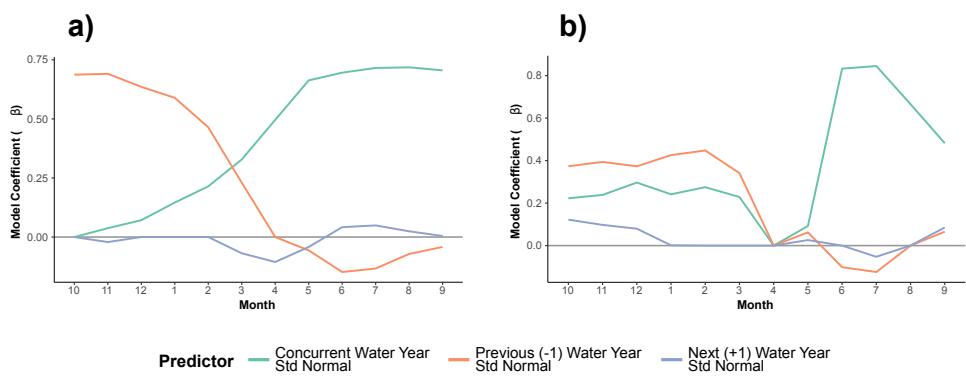


Figure 3: APR model coefficients for concurrent and lagged (-1, +1) years at the (a) Logan and the (b) Bear rivers.

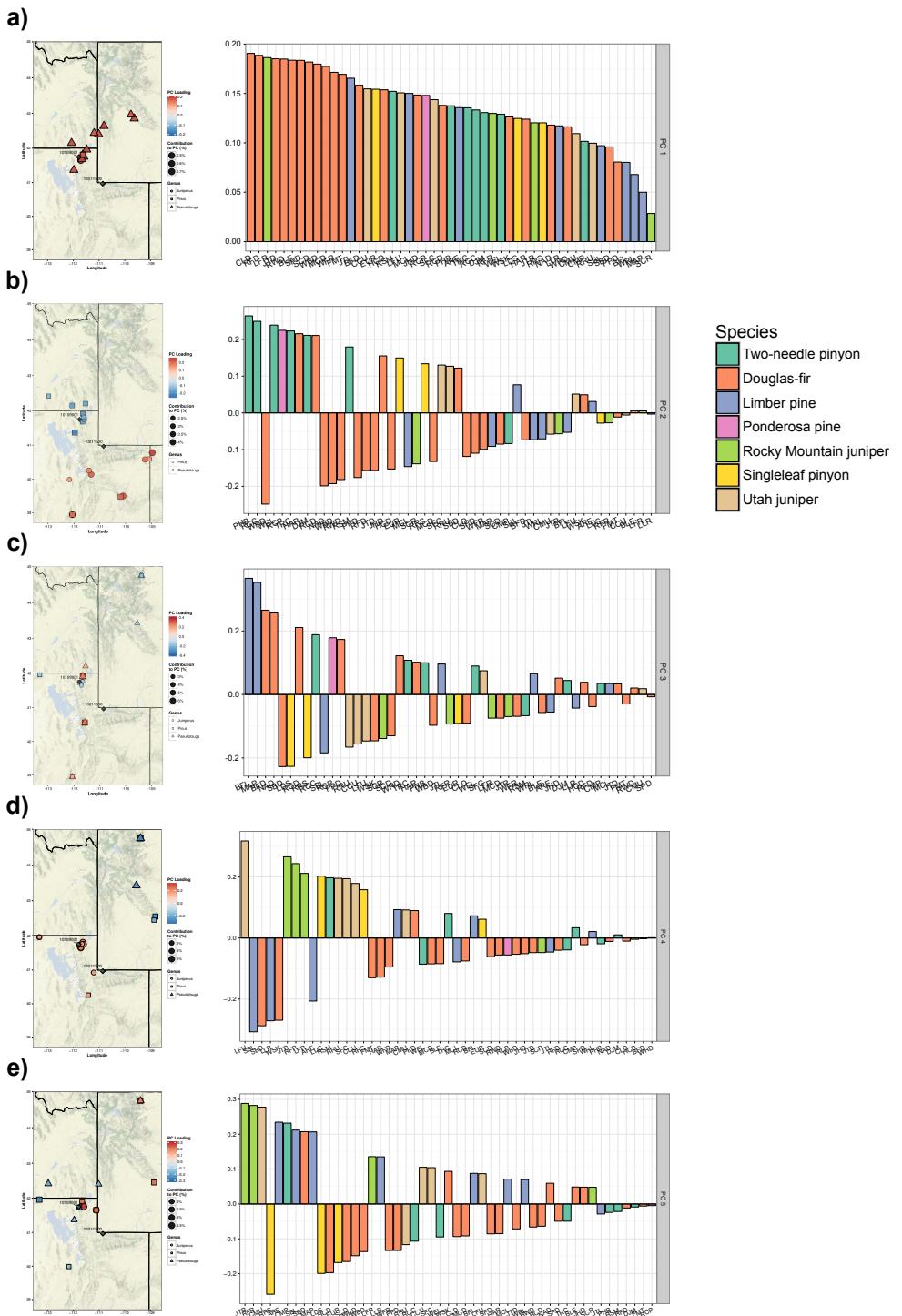


Figure 4: PC loadings from imputed values presented spatially (left) and by species (right). Loadings are shown for PCs 1-5 (a-e). Spatial maps only present sites with loadings ≥ 0.20 or than the mean to prevent over-plotting.

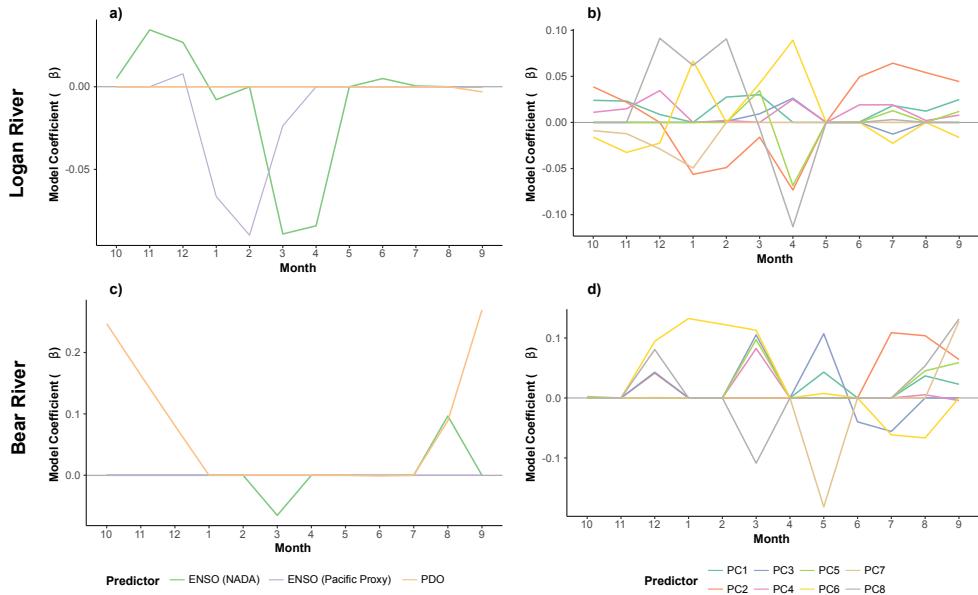


Figure 5: Model coefficients for the Logan River (a-b) and the Bear River (c-d). Global circulation indices are shown on the left (a,c), while tree-ring PCs are shown on the right (b,d).

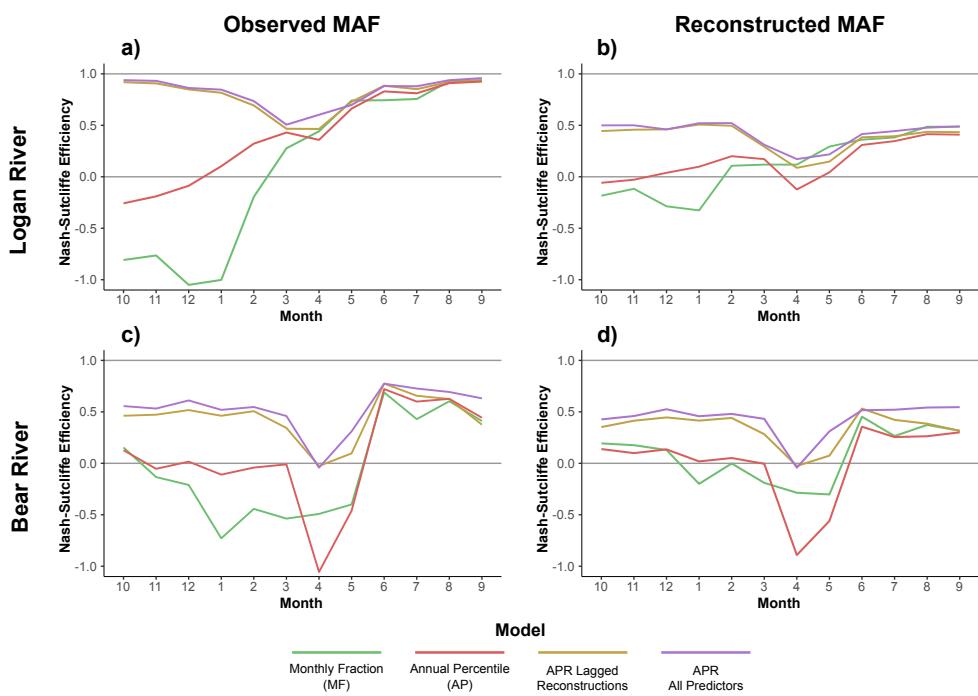


Figure 6: Nash-Sutcliffe Efficiency (NSE) by month.

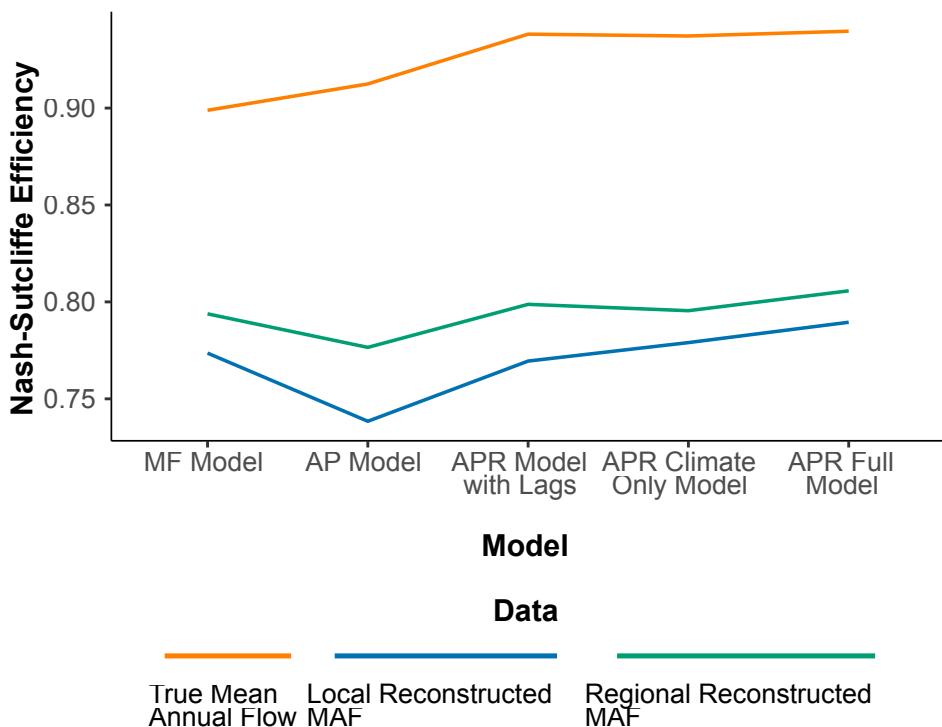


Figure 7: Nash-Sutcliffe Efficiency (NSE) for the Logan River. Models of increasing complexity are presented along the horizontal axis, while colors represent the input reconstruction dataset.

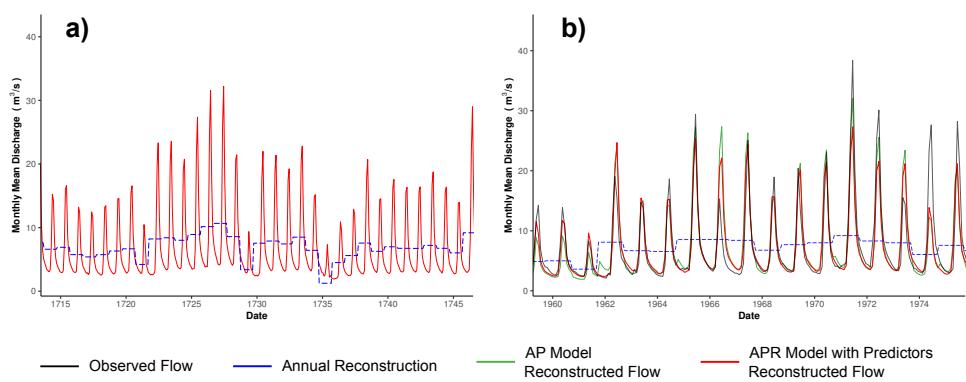


Figure 8: Reconstructed flows at the Logan River site for subsets of the (a) historical and (b) observed periods.

690 **List of Tables**

691	1	Annual goodness of fit for Logan River. ME represents mean error, MAE repre-	
692		sents mean absolute error, RMSE represents root mean squared error, NSE rep-	
693		resents Nash-Sutcliffe Efficiency, R represents Pearson correlation, and $R_{Spearman}$	
694		represents Spearman's rank correlation.	35

Table 1: Annual goodness of fit for Logan River. ME represents mean error, MAE represents mean absolute error, RMSE represents root mean squared error, NSE represents Nash-Sutcliffe Efficiency, R represents Pearson correlation, and R_{Spearman} represents Spearman's rank correlation.

Model	MAF	ME (m^3/s)	MAE (m^3/s)	RMSE (m^3/s)	NSE	R	R_{Spearman}
MF	Observed	0.012	1.25	1.97	0.899	0.951	0.892
AP	Observed	0.083	1.11	1.83	0.912	0.958	0.900
APR + Lags	Observed	-0.069	0.821	1.54	0.938	0.969	0.971
APR + Predictors	Observed	-0.099	0.803	1.51	0.940	0.970	0.975
MF	Reconstructed	0.065	1.71	2.94	0.774	0.881	0.872
AP	Reconstructed	-0.054	1.78	3.16	0.738	0.868	0.853
APR + Lags	Reconstructed	-0.190	1.60	2.97	0.769	0.879	0.899
APR + Predictors	Reconstructed	-0.135	1.53	2.82	0.790	0.889	0.913