AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2020WR027706

Key Points:

- Tree growth across the Mid-Atlantic is more closely related to the baseflow component of streamflow than to total streamflow
- Stormflow can represent noise in paleoclimatic reconstruction exercises and impart increased uncertainty for preinstrumental estimates
- Two seasonally distinguished reconstructions of Potomac River streamflow are presented for the past 350+ years

Supporting Information:

- Supporting Information S1

Correspondence to:
M. C. A. Torbenson,
torbenson.1@osu.edu

Citation:
Torbenson, M. C. A., & Stagge, J. H. (2021). Informing seasonal proxy-based flow reconstructions using baseflow separation: An example from the Potomac River, United States. *Water Resources Research*, *57*, e2020WR027706. https://doi.org/10.1029/2020WR027706

Received 10 APR 2020
Accepted 22 NOV 2020

Informing Seasonal Proxy-Based Flow Reconstructions Using Baseflow Separation: An Example From the Potomac River, United States

M. C. A. Torbenson[1] **and J. H. Stagge**[1]

[1]Department of Civil, Environmental and Geodetic Engineering, Ohio State University, Columbus, OH, USA

Abstract Paleoclimatic perspectives on hydrological variability can offer valuable information on the frequency and magnitude of extreme events. Tree-ring records are a common proxy used to reconstruct past streamflow due to their interannual resolution and often strong correlation with hydroclimate variability. The separation of streamflow into theoretical baseflow and stormflow constituents is regularly utilized to differentiate between flow-generating processes; however, it has yet to see use in paleoclimate reconstructions. We compare three approaches which use log-linear, gamma-distributed, and baseflow-separated regression, respectively, to reconstructing flow at a well-studied gage—the Potomac River at Point of Rock, Maryland. Preinstrumental baseflow and streamflow for summer were estimated for the past 350 years using a regional network of tree-ring chronologies. Additionally, estimates of winter flow were produced for the same period using a nonoverlapping set of tree-ring data. Tree growth appears to have a stronger relationship with baseflow than with streamflow for both seasons, supporting the use of baseflow as predictand. The number of chronologies subsequently chosen as predictors for streamflow was also lower than for baseflow and represents an additional source of reconstruction bias/uncertainty when total streamflow is the predictand. Historical records support the validity of both summer and winter reconstructions. The winter reconstruction indicates that several years of consecutive below-mean flows, on par with the 1960s drought, occurred with higher frequency prior to the instrumental era. Our results suggest that baseflow separation can improve reconstruction skill and provide additional information to water resource management on the long-term variability of hydrological extremes.

1. Introduction

Estimating return periods for extreme low-flow and flood events may result in large associated uncertainties due to the relatively short observational records of streamflow (Koutsoyiannis & Montanari 2007), especially in regions where climatological influence on streamflow variability has a low-frequency component (e.g., Enfield et al., 2001). Paleoclimate reconstructions based on tree-ring chronologies have been used to extend instrumental streamflow data across the world (e.g., Gou et al., 2007; Meko et al., 2020; Urrutia et al., 2011; Rao et al., 2018). Notable reconstructions from the United States include past estimates for the San Joaquin-Sacramento River (Meko et al., 2001) and the Colorado River (e.g., Woodhouse et al., 2006). These types of proxy records play a potentially increasing role in U.S. water management for understanding long-term variability (Rice et al., 2009; Tingstad et al., 2014; Woodhouse et al., 2016), runoff efficiency (Cleaveland & Stahle, 1989; Woodhouse & Pederson, 2018), as well as the frequency, magnitude, and persistence of extreme events (Gonzales & Valdes, 2003; Razavi et al., 2015; Woodhouse et al., 2013).

Traditional dendroclimatic studies have focused on identifying the strongest climate signal in the tree-ring record (Cook et al., 1999), often an aggregated soil moisture variable such as growing-season Palmer Drought Severity Index (PDSI, Palmer, 1965). However, this aggregate signal is driven by different seasonal precipitation components depending on species and local climatology (St. George et al., 2010; Stahle et al., 2020). In recent decades, an increasing number of seasonally resolved hydroclimate reconstructions have been produced (e.g., Stahle et al., 2009; Torbenson & Stahle, 2018). The extrapolation of warm season streamflow signal in tree rings has also been shown to have skill on bi-weekly to monthly scales (Sauchyn & Ilich, 2017; Stagge et al., 2018). Although significant work has been done to improve our understanding of the predictor variables of tree-ring based streamflow reconstruction, including predictor selection (Ho et al., 2016; Saito et al., 2008; Strange et al., 2019), standardization techniques (Meko et al., 2015), and the

application of Bayesian modeling (Devineni et al., 2013; Ravindranath et al., 2019), less attention has been given to the target variable to be reconstructed—the instrumental streamflow data.

The Potomac River is the main source of water for the Washington, DC, metropolitan area. Increases in the region's population are expected for the 21st century, which will be coupled with net increases in water usage (Bhatkoti et al., 2018; Schultz et al. 2017). Forecasts estimate a 12% increase in average annual water demand between 2015 and 2040 (Ahmed et al., 2015). Water vulnerability, especially due to drought, is a real concern and long-term estimates can therefore provide a greater understanding of the range of conditions that may be experienced in a changing climate (Stagge & Moglen, 2017; Tebaldi et al., 2006). Three reconstructions of Potomac River summer streamflow have previously been produced using tree-ring data (Cook & Jacoby, 1983; Maxwell et al., 2011, 2017). These reconstructions suggest that there have been years and periods in the past for which the trees indicate conditions outside the range of observational records for the of the 20th and early 21st centuries, from which the "drought of record" is defined for planning purposes.

In this paper, we revisit the relationship between eastern U.S. tree growth and Potomac River streamflow variability. Using daily instrumental streamflow data at Point of Rock, MD, baseflow and stormflow components are separated, accumulated to seasonal means, and ultimately correlated to members of a network of regional tree-ring chronologies. Three reconstruction approaches are compared: (1) a traditional linear model using a log-transformed streamflow reconstruction target; (2) a generalized linear model (GLM) with linked gamma-distributed streamflow for a target; and (3) a GLM with a gamma-distributed baseflow target. Reconstructions of June-to-August (JJA) base and streamflow are produced spanning 1666 to the present day. In addition, experimental reconstructions of December-to-February (DJF) base and streamflow are presented for the same period.

2. Data and Methods

To test the differing roles of theoretical flow constituents in tree growth signals, and the effect of this choice on subsequent reconstructions, tree-ring chronologies from the Mid-Atlantic of the eastern US were correlated with instrumental Potomac River streamflow data, as well as separated baseflow and stormflow. The three candidate reconstruction models were compared for a common calibration period for cool and warm seasons, resulting in a total of six models.

2.1. Instrumental Streamflow

Daily instrumental streamflow data were based on the United States Geological Survey (USGS) gage 01638500 at Point of Rocks, MD (39°16′24.9″N, 77°32′35″W). The full 01638500 gage series begins in 1895 and extends to the present. The instrumental gage data was adjusted by the Interstate Commission on the Potomac River Basin (ICPRB) to create a naturalized flow series by reinserting upstream withdrawals and reservoir impacts. The resultant naturalized series covers the period October 1, 1929 to November 30, 2007. While this limits the available length for model calibration purposes, it more accurately represents the river in its natural state, as would be captured by tree-ring proxies, and avoids two major streamflow gage changes in September 1902 and October 1929. Beginning the model calibration after 1929 also avoids some of the most severe clearcutting of forests during the 1895–1910 period, which might also introduce inhomogeneities or trends into the underlying flow signal (Cook & Jacoby, 1983). The original, un-naturalized instrumental series was used for validation purposes between 1895 and 1929, as it provides a convenient hold-out period with gaged data. Direct comparison of reconstructed naturalized flow and gaged un-naturalized flow is reasonable for this validation period because it is prior to construction of the Savage and Jennings Randolph Reservoirs, in 1953 and 1981, respectively, making naturalization relatively insignificant.

The Potomac River watershed covers 38,000 km^2 and four states, with its headwaters in eastern West Virginia and western Virginia/Maryland. Streamflow at the Point of Rocks gage historically peaks in March and reaches its lowest levels during late summer (July, August, and September). Point of Rocks is a critical gauge for water management decisions in the Washington, DC, metropolitan area because it lies 2.4 days upstream of the Washington DC water treatment intakes (Ahmed et al., 2015; Hagen et al., 2005). This 2.4 days buffer allows water managers to forecast flow at the operational target Little Falls gauge using a simple flow accumulation model. Managers can then anticipate shortages during low-flow periods and fine-tune

reservoir release or load shifting decisions. The Potomac River is located in a transitional region for the observed effects of climate change, with flows generally predicted to decrease to the south and increase to the north of the watershed (Yang et al., 2015). Observed and projected changes in the Potomac have been more nuanced and seasonal, with increases in winter/early spring flow and decreases in summer flow (Rice & Hirsch, 2012; Stagge & Moglen, 2013). These trends are not large relative to annual vulnerability, providing confidence in the period for use in proxy-calibration.

2.2. Baseflow Separation

Separation of streamflow into baseflow and stormflow components was based on separating the daily instrumental time series, which was then aggregated to monthly baseflow and stormflow for reconstruction purposes. Baseflow separation was performed by the HYSEP sliding interval method (Sloto & Crouse, 1996) using the DVstats R package developed by the United States Geological Survey (USGS, Lorenz, 2017). The sliding interval method is an empirical separation approach that estimates baseflow using the minimum flow during a period centered on the day in question. The width of the moving window period is typically chosen as twice the travel time for storm runoff ($2N$) rounded to the nearest odd integer, where N represents the number of days after which surface runoff effects cease following a storm event (Pettyjohn & Henning, 1979). For this study, we assumed the empirical relationship provided in HYSEP guidance of $N = 0.83 \times A^{0.2}$ where A is the drainage area in square kilometers, resulting in N of 6.5 days and a sliding interval width of 13 days. This matches well with traveltime observations during low-flow operations, which showed a traveltime of approximately 9–10 days from the headwaters to Little Falls (Ahmed et al., 2015). During low-flow conditions, lag between Point of Rocks (reconstruction target) and the downstream Little Falls is 2.4 days, making travel time to Point of Rocks approximately 6.6–7.6 days. Lag times vary with flow, but this 6.6–7.6 days range agree with our proposed sliding interval of 6.5 days.

The HYSEP sliding interval method was chosen because of its common use within the hydrologic community, simplicity of interpretation, and performance in prior studies (Eckhardt, 2008; Gonzales et al., 2009; Partington et al., 2012; Schultz et al. 2014). Identifying the true baseflow contribution in observations is difficult without involved tracer studies, and even then, it is uncertain. However, in simulated small watersheds the sliding interval outperformed other separation techniques for sandy soils, though underperformed in sandy loam soils (Partington et al., 2012), though it underperforms in sandy loam soils. It should be noted that this simulated watershed was four orders of magnitude smaller than the Potomac. The Potomac watershed has sufficient elevation gradient, which avoids the documented issue of poor performance in flat, lowland watersheds (Gonzales et al., 2009). Several other approaches to flow separation exist (Lott & Stewart, 2016). To further confirm our choice of baseflow separation method, we compared the resultant baseflow timeseries with an alternative UKIH method (Piggott et al., 2005). The differences between the resulting timeseries were minor for the 1931–1980 period ($r = 0.987$ and 0.992 for DJF base and stormflow respectively; 0.951 and 0.996 for JJA). Results presented for the remainder of this study are based solely on the HYSEP sliding interval series.

2.3. Tree-Ring Chronologies

Earlywood (EW), latewood (LW), and total-ring width (TRW) chronologies from the northeastern United States (32–48°N, 65–87°W) available through the International Tree-Ring Data Bank (ITRDB; Zhao et al., 2019) were correlated with monthly resolved baseflow, stormflow, and streamflow series, as well as with seasonalized flow totals for DJF and JJA, for the full common period 1931–1980. In most dendroclimatological reconstructions utilizing a large network of data, only a subset of chronologies is used as predictors for the final estimates of past variability because of differing climatic signals (e.g., Cook et al., 1999). Because the candidate reconstruction models have different target variables (total streamflow, baseflow), these models could dictate different sets of predictor chronologies. To ensure a fair model comparison, we apply a common correlation threshold to objectively select a subset of chronologies tailored for each model. Chronologies that displayed a Spearman correlation >0.37 ($p < 0.01$) with seasonally resolved flow were selected as possible time series for Principal Components Analysis (PCA; Jolliffe, 2002). The final predictor pools contain EW, LW, and TRW records. To minimize exaggeration of

season-to-season persistence, no TRW chronology could be considered for both cool and warm season (Stahle et al., 2020; Torbenson & Stahle, 2018). If a chronology was significantly correlated with aggregates of both seasons, it was considered a predictor variable for the season in which it displayed the highest correlation. For EW and LW records, alternate chronologies from the same site were allowed as predictors for different seasons. For the JJA window, chronologies with no significant correlations with DJF or JJA flow for year i were considered as lagged predictors (flow in year i_{-1}) as some tree species are known to store photosynthate (e.g., Aloni, 1991).

2.4. Candidate Reconstruction Models

The first PC in year i from this analysis served as the predictor in three different reconstruction approaches:

(1) A simple linear regression model (M1) with log-transformed streamflow as the predictand

$$\log\left(\hat{Q}_{stream,i}\right) = \beta_0 + \beta_1 PC_{1,i}$$

(2) A generalized linear model (M2) with streamflow fitted using a gamma distribution with shared shape parameter, υ, and a scale parameter, θ_i, which in turn is controlled by the identity link with the expected value, g (μ_i)

$$\hat{Q}_{stream,i} \sim \text{Gamma}\left(\upsilon, \theta_i\right), \text{where } \theta_i = \mu_i / \upsilon \text{ and}$$
$$g\left(\mu_i\right) = \beta_0 + \beta_1 PC_{1,i}$$

(3) A generalized linear model (M3) with baseflow fitted using a gamma distribution

$$\hat{Q}_{base,i} \sim \text{Gamma}\left(\upsilon, \theta_i\right), \text{where } \theta_i = \mu_i / \upsilon \text{ and}$$
$$g\left(\mu_i\right) = \beta_0 + \beta_1 PC_{1,i}$$

The stormflow component is estimated and added to the reconstruction to produce total streamflow estimates (see below).

Model 1 was performed to establish a baseline, mimicking the regression approaches used in prior studies (Cook & Jacoby, 1977; Maxwell et al., 2011, 2017). Model 2 uses a Generalized Linear Model (GLM) with gamma-distributed response through the identity link function, which assumes that the residuals of the target predictand are gamma distributed with a common shape parameter (1/dispersion) and a scale parameter that varies with the estimate. It is hypothesized that because flow is strictly positive (>0) and tends to have a positive skew, with a long upper tail, a gamma-distributed response will produce a reconstruction with less bias. Further, predictions from a regression model fit in log-space and backtransformed can be systematically biased because of Jensen's Inequality (Ruel & Ayres, 1999). Ignoring this violation of regression assumptions can over-emphasize high flows in proxy-calibration. Model 3 uses the same Gamma GLM framework, but instead seeks to reconstruct the baseflow component from the proxy and estimate the remaining stormflow postcalibration. The stormflow estimates for M3 were calculated by adding the portion of variability shared with baseflow each year and the mean climatology of the independent stormflow (based on the instrumental data):

$$\hat{Q}_{stream} = \hat{Q}_{base} + \hat{Q}_{storm}$$
$$\hat{Q}_{storm} = K \cdot \left[\hat{Q}_{base} \cdot r^2{}_s + \bar{Q}_{storm} \times \left(1 - r_s^2\right)\right]$$

where $\hat{Q}$ represents estimated for each of the three components: streamflow, stormflow, and baseflow. $\bar{Q}_{storm}$ represents the mean climatology of stormflow, $r^2{}_s$ represents $r^2\left(Q_{storm}, Q_{base}\right)$, the historical percent shared variance of stormflow explained by baseflow. K represents the scale factor of the relative contribution of

stormflow in streamflow. The uncertainty estimates for M3 streamflow include the range of the independent climatology of stormflow, relative to the total contribution of stormflow.

2.5. Model Calibration and Validation

Model calibration was performed on the full common period 1931–1980 to limit the risk of overfitting and any possible bias from extended periods of persistent conditions (e.g., the 1960s drought). Because of differing start dates of the predictor chronologies, a nested approach in which the number of tree-ring chronologies decreased back in time was taken. The nests were spliced together, using the nest of maximum number of chronologies entered into the PCA for each segment to produce a continuous time series. Reconstruction later than 1980 was not feasible because several chronologies are not available for the more recent period.

The resulting reconstructions were compared with pre-1930s instrumental streamflow data as a form of validation. As described above, these early observations have not been naturalized and may have been influenced by human activity (Cook & Jacoby, 1983), but human effects should be relatively small, allowing for further verification of the models, outside the calibration period. Furthermore, the naturalized time series represents the interannual variability of un-naturalized flow, and therefore should not affect any correlation analysis. In addition, selected years of reconstructed extreme flows were compared to historically documented records of drought and flood.

3. Results

3.1. Proxy Correlations

A total of 193 TRW chronologies and 25 pairs of EW/LW chronologies were screened for correlation with Potomac River flow variability. The number of chronologies that display significant correlations ($p < 0.01$) with total stormflow (Figure 1c) is lower than for both total streamflow and baseflow (Figures 1a and 1b, respectively). Two distinct seasons of significant correlations are visible for stream and baseflow: December-February and June-August (Figure 1). Few chronologies show significant correlations with March-May flow; the transitional period between these seasons and also the period of peak annual flow at Point of Rocks. The number of chronologies showing significant correlations with mean flow for the cool (DJF) and warm (JJA) season windows are presented in Table 1.

The locations of selected tree-ring chronologies (Figure 2; Tables S1 and S2) largely agree with the correlation between seasonal streamflow and early/late growing-season soil moisture and precipitation (Figure S1). No EW/LW pairs were selected for both seasonal predictor pools. The positive relationship between DJF streamflow and May PDSI extends from the Potomac watershed to the Canadian border and as far west as Missouri. Correlations between summer precipitation and streamflow are less extensive, and mainly confined to the basin. Counts are lower for streamflow than for baseflow for both seasons, but the difference is greater for JJA. The loss of chronologies due to the discrete condition (predictors cannot be shared between seasons) is greatest for the JJA streamflow reconstructions (M1/M2) and the DJF baseflow model (M3). Overall, the average correlation between selected chronologies and flow does not differ significantly between models/selection criteria.

Correlations between the reconstructed values of streamflow and the un-naturalized early instrumental data (1897–1930) are generally low for both seasons (Table 2). The only significant correlation for the validation period present in the models is exhibited by M3 for JJA ($r = 0.527$). Although the precalibration validation statistics only suggest modest skill in the reconstructions, comparisons with contemporaneous historical records corroborates many years of extreme flow prior to 1931 (see below).

3.2. Model Calibration and Validation

The DJF M1 explains 36% of the variance in instrumental streamflow for the calibration period (Figure 3a). The use of a GLM and a fitted gamma distribution (M2) produces marginally better statistics (39% explained

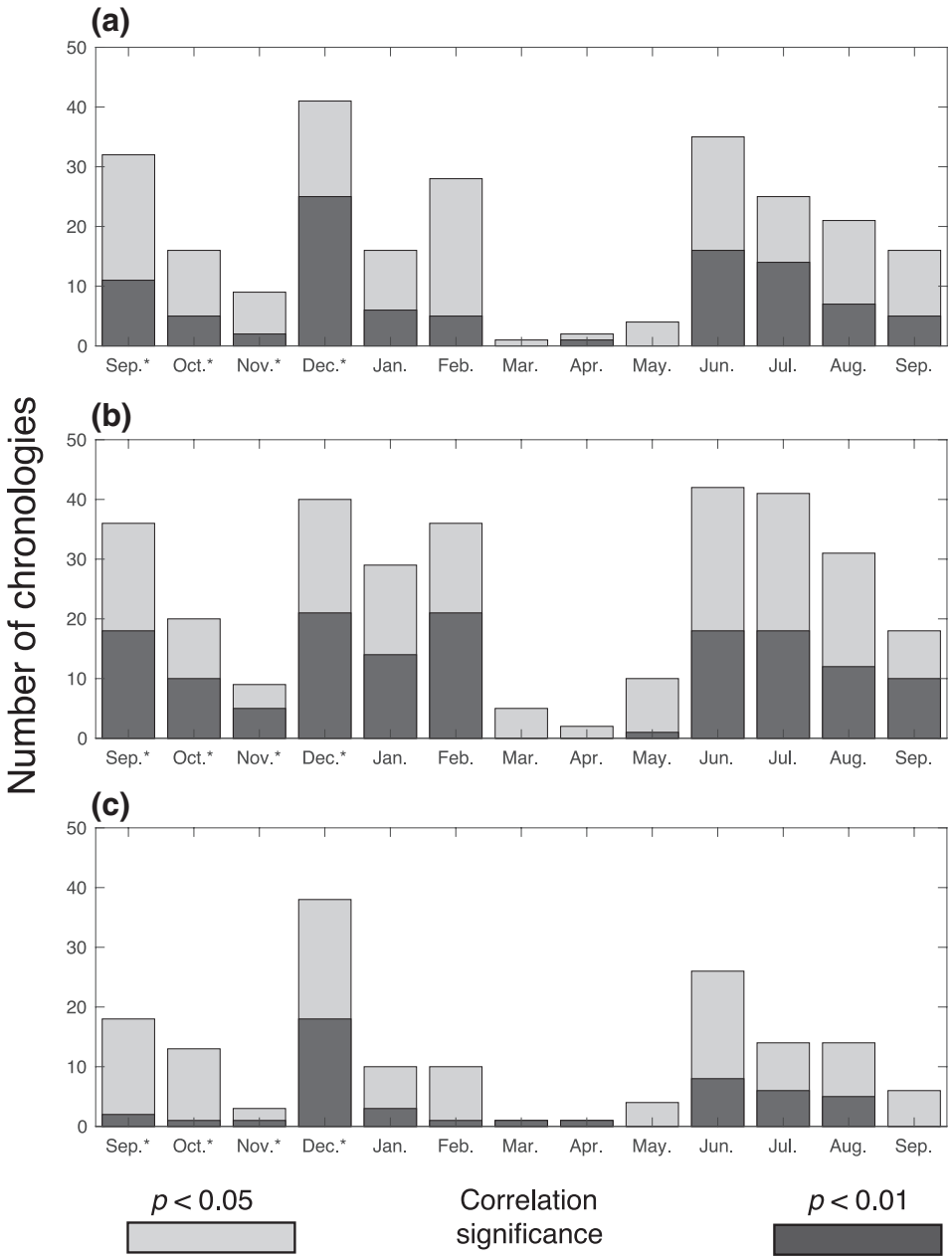

Figure 1. Number of tree-ring chronologies significantly correlated ($p < 0.05/0.01$) with monthly (a) streamflow, (b) baseflow, and (c) stormflow at Point of Rocks, MD, for the period 1931–1980. Asterisks denote months of the previous year.

variance). The third model explains 43% of the baseflow variance but when adding the stormflow component of streamflow to the final reconstruction the statistics weaken to 37%. For JJA (Figure 3b), M1 displays less skill than M2 and M3 but the overall results are weaker than for DJF, ranging from 24% to 26% of the explained variance. It is worth noting that the baseflow reconstruction of M3 for JJA explains 44% of the variance during the calibration period, suggesting that the hydroclimate signal in tree growth is more attuned to the baseflow constituent in summer than during the start of the growing season.

Due to the relatively weak overall relationship between instrumental and reconstructed data, the reconstruction variances are lower than the instrumental data for the calibration period. The loss of explained variance is greater for JJA than for DJF. The most replicated nests for M3 start in 1755 (DJF) and 1750 (JJA),

Table 1
Counts of Chronologies, and Their Average Correlation, That Display Significant Correlation With Base/Streamflow for Each Season and Were Used as Predictors

	# Chrons (DJF_dis)	Ave. Corr.	# Chrons (DJF_all)	Ave. Corr.
M1/M2	16	0.442	19	0.441
M3	17	0.453	26	0.451
	# Chrons (JJA_dis)		# Chrons (JJA_all)	
M1/M2	10	0.429	17	0.433
M3	18	0.458	23	0.467

DJF_dis and JJA_dis indicate counts of discrete (i.e., nonoverlapping) chronologies only and DJF_all and JJA_all indicate nondiscrete counts.

and the resulting reconstructions span 1666 to 1980 (Figure 4; Figure S2; statistics of the individual nests are given in Table S3).

3.3. Residual Analysis

The reconstructed DJF baseflow is significantly correlated with instrumental DJF stormflow ($r = 0.498$) for the calibration period, likely due to the natural hydrologic correlation between baseflow and stormflow captured in the instrumental record ($r = 0.777$; Figure 5). To assess the dominant signals in the tree-ring records, residuals from a simple regression of the M3 baseflow reconstruction ($\hat{Q}_{base}$) on the instrumental stormflow (Q_{storm}; Figure 5a) and the instrumental baseflow (Q_{base}) on the instrumental stormflow (Q_{storm}; Figure 5b) were produced. The residuals from these two regressions are significantly correlated over the calibration period ($r = 0.486$; Figure 5c). Regressing the instrumental stormflow on the reconstructed baseflow, and the instrumental stormflow on the instrumental baseflow, produces two residual time series that are not correlated ($r = -0.017$; not shown), suggesting that the baseflow reconstruction does not contain any "leftover" stormflow information.

Similarly, the estimated JJA baseflow is positively correlated with instrumental JJA stormflow ($r = 0.369$), although significantly weaker than the instrumental base and stormflow correlation ($r = 0.726$; Figure S3ab). The difference in skill between baseflow and streamflow models is greater for the JJA seasonal window (Table 2). Correlation between the residual time series is high ($r = 0.618$; Figure S3c). The residual

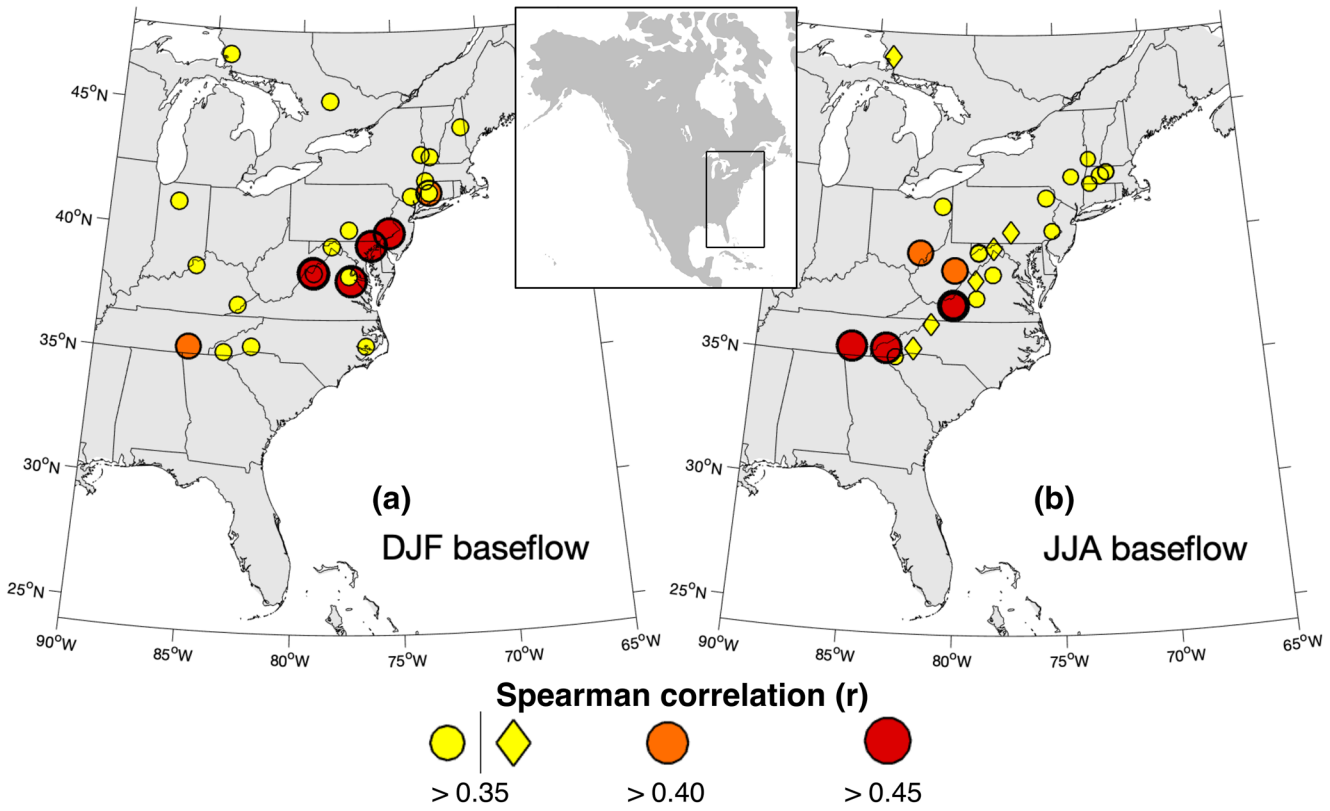

Figure 2. Location of tree-ring chronologies displaying significant correlations ($p < 0.01$) with (a) DJF and (b) JJA baseflow at Point of Rocks, MD, for the period 1931–1980. Circles represent predictors from year i, diamonds represent predictors from year $i + 1$ (JJA only). Black outline in map inset indicates location of region in (a) and (b). DJF, December-to-February; JJA, June-to-August.

Table 2
Calibration Statistics (1931–1980) and Correlation With Precalibration Instrumental Streamflow Data for the Three Models

	Calibration		Validation
DJF	Base (r^2)	Stream (r^2)	USGS 1897 – 1930 (r)
M1	N/A	0.364	0.184
M2	N/A	0.387	0.179
M3	0.426	0.370	0.187
JJA	Base (r^2)	Stream (r^2)	USGS 1897 – 1930 (r)
M1	N/A	0.239	0.216
M2	N/A	0.258	0.244
M3	0.440	0.248	0.527

"independent" stormflow correlation for JJA is negative but not significant ($r = -0.216$). JJA stormflow and baseflow are more subject to high outliers than the winter season.

4. Discussion

Reconstructing streamflow from tree-ring chronologies comes with numerous challenges. Streamflow for many U.S. rivers does not follow a normal distribution, as is the case for the Potomac River at Point of Rocks. Furthermore, streamflow variability is the result of several different hydrological processes that may not evenly be connected to tree growth. Our results suggest that instrumental data treatment, such as the separation of flow into base and storm constituents, and alternative regression approaches (including the application of GLMs) can mitigate some of these issues and help provide more accurate estimates of reconstruction uncertainties.

4.1. Model Comparison

The skewed distribution of streamflow can cause negative reconstructed values if using raw data as the predictand in a traditional simple regression model that assumes gaussian residuals, including near zero. Maxwell et al. (2011, 2017) used a log-transformed target to address this issue, however, the resulting reconstruction underestimates the mean summer streamflow for the instrumental era (similar to Cook & Jacoby, 1983). Part of this underestimation may stem from Jensen's Inequality (Ruel & Ayres, 1999), which describes how backtransforming can consistently over-predict or under-predict in the original coordinates. Although the year-to-year variability still represents the streamflow signal in tree growth, the log-transformation adds additional obstacles when comparing instrumental data with reconstruction estimates.

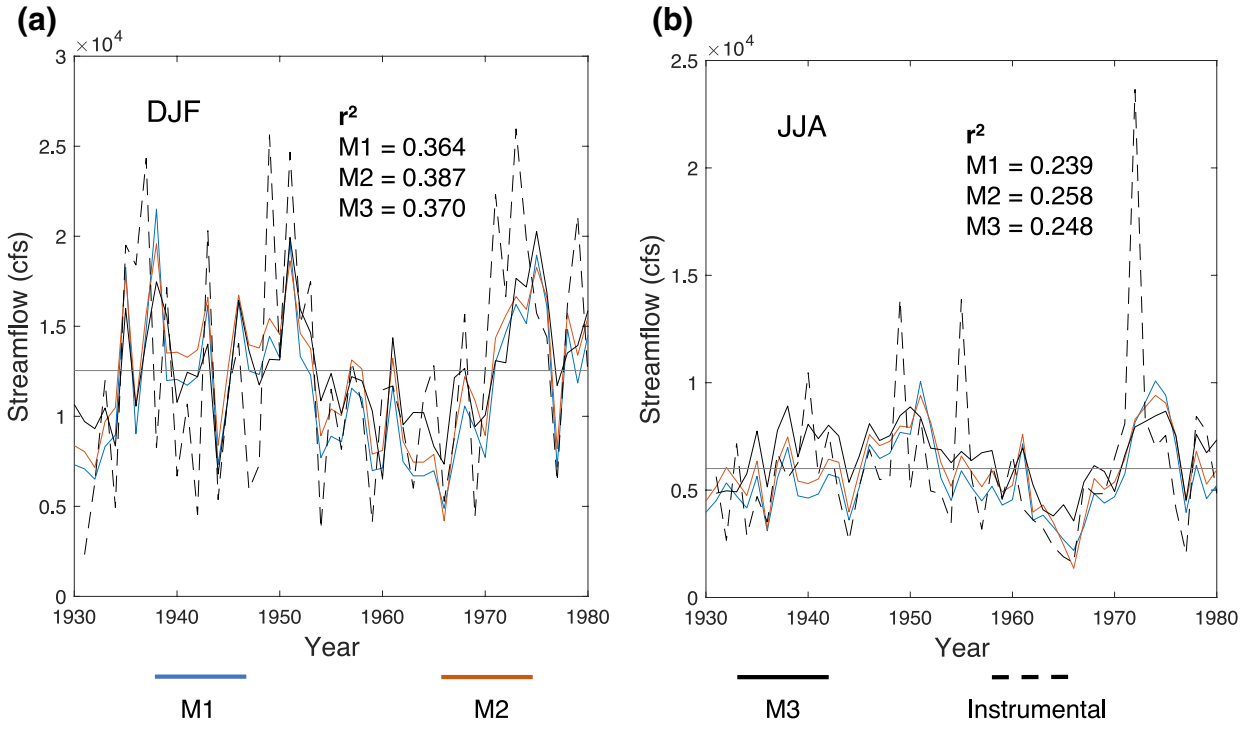

Figure 3. Time series comparison between instrumental (dashed line) and three different models of reconstructed streamflow for (a) DJF and (b) JJA. The gray horizontal line represents the instrumental mean for 1931–1980. DJF, December-to-February; JJA, June-to-August.

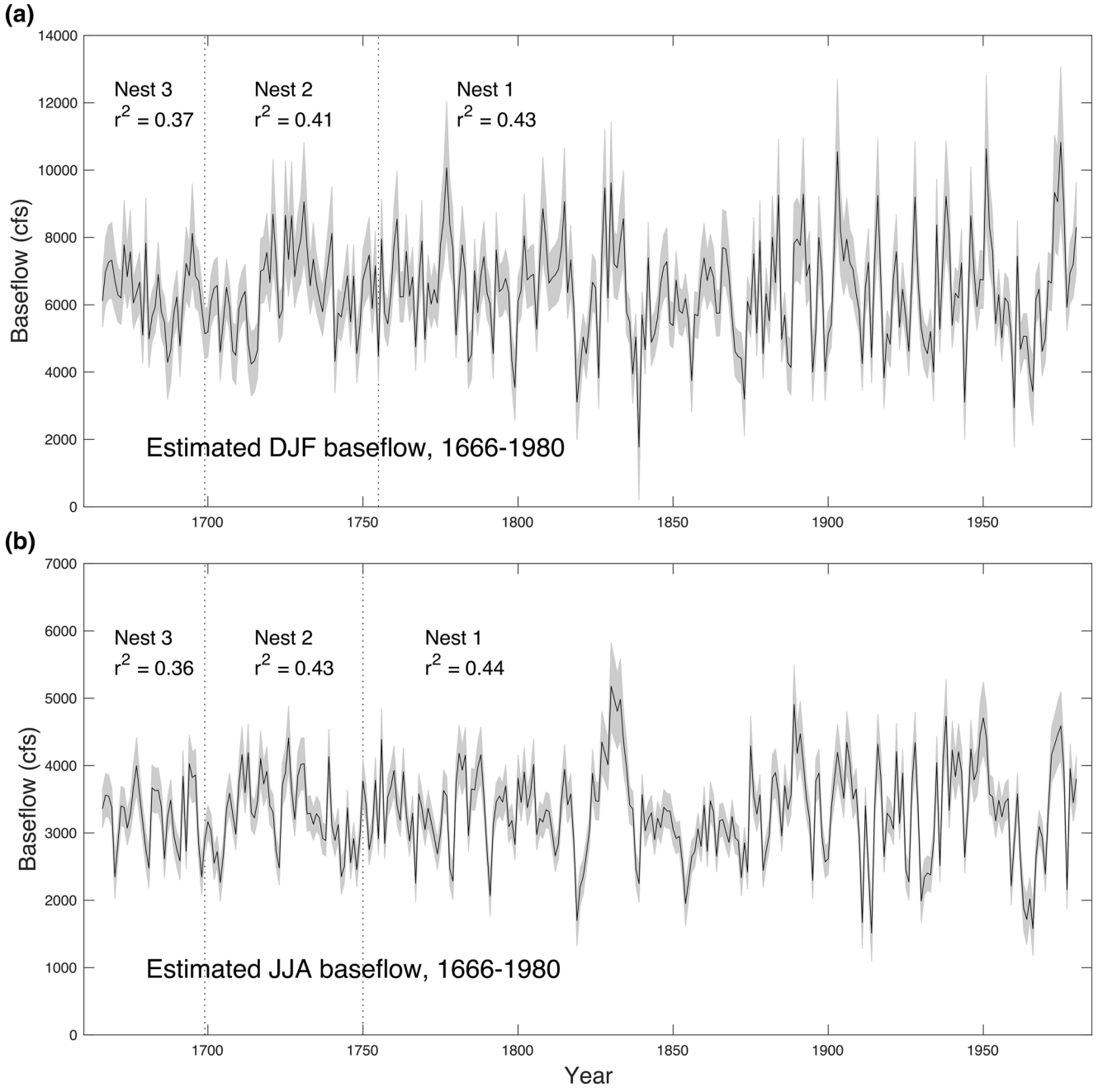

Figure 4. Model 3 reconstructions of baseflow at Point of Rocks, MD, extending from 1666 to 1980 based on two separate networks of tree-ring chronologies in the Mid-Atlantic region: (a) DJF and (b) JJA, with the gray band representing 5–95% prediction confidence interval. DJF, December-to-February; JJA, June-to-August.

Reconstruction statistics calculated for the log-space do not necessarily represent the relationship between instrumental flows and estimated data in true streamflow units (Prairie et al., 2006) and therefore risk overstating the skill of the reconstruction. Such biases can be accounted for (e.g., Helsel et al. 2020) but are for comparison purposes not considered here.

The GLM approach produces slightly stronger statistics than the linear regression with a log-transformed target for the calibration period (Table 2) but the differences are not statistically significant for either season ($p > 0.05$; Fisher, 1921). All models, including the linear log-transformation (M1), capture the mean flow

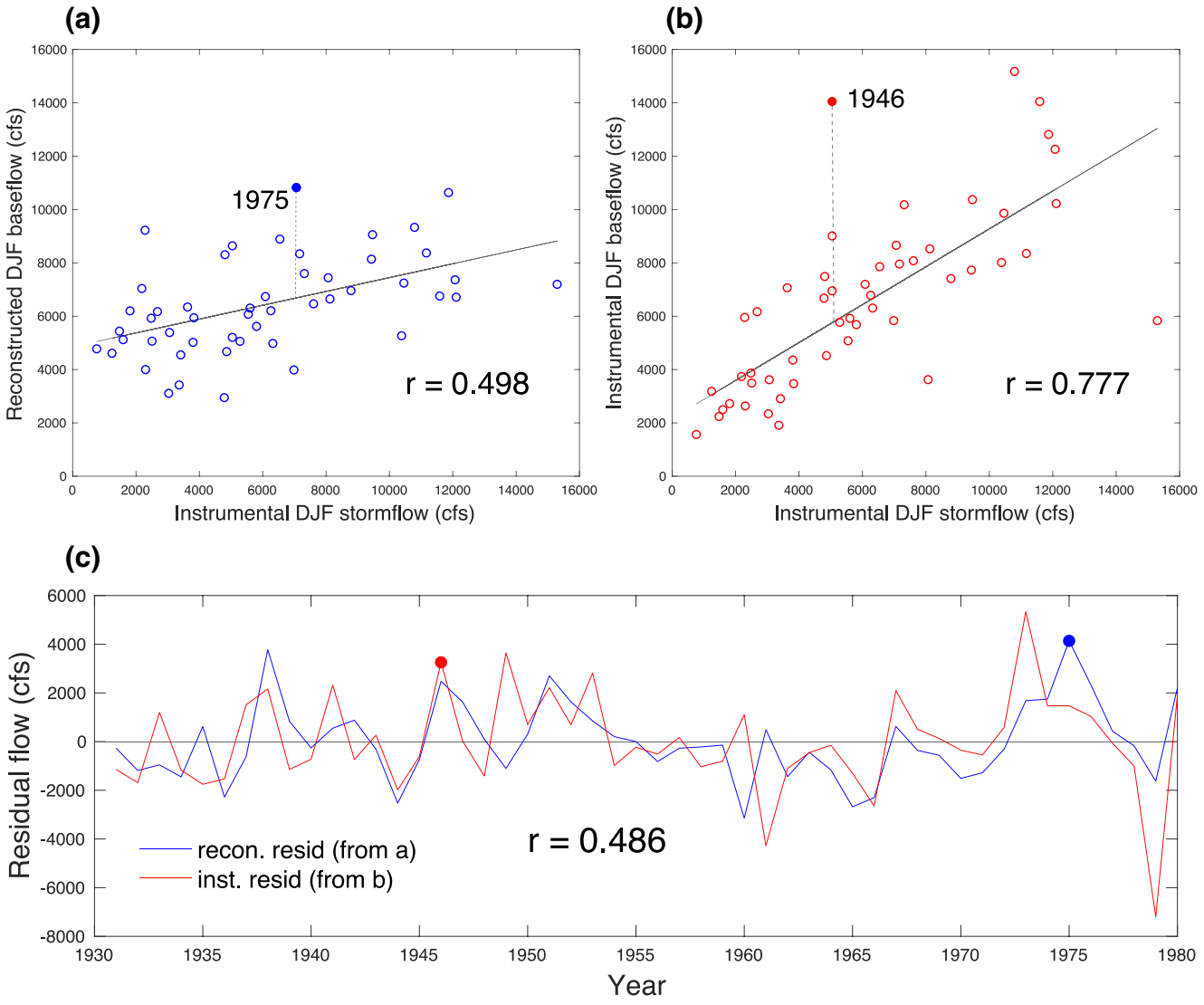

Figure 5. Scatterplot between reconstructed DJF baseflow and instrumental DJF stormflow (a), and instrumental DJF baseflow and DJF stormflow (b), for the period 1931–1980. A time series comparison between the residuals of a simple regression (panels a and b) is presented (c). DJF, December-to-February.

of the instrumental data for the calibration period, which is expected with simple one-predictor regression models. No significant differences between M1 and M2 validation correlations are recorded. Because the two models utilize the same predictor set and are calibrated on the same target, these results must be considered predictable. Model three is discussed in detail below.

4.2. Streamflow Versus Baseflow

The results from correlating tree-ring records with separated constituents of flow suggest that the interannual streamflow signal exhibited in trees around the Potomac River basin is predominantly connected to baseflow. The variability of growth appears to be less correlated with direct overland runoff, evident by the low number of chronologies significantly correlated with stormflow (Figure 1). This is despite stormflow making up a significant part of total streamflow volume and being closely tied to baseflow variability. This difference follows logically from the catchment perspective, in which precipitation is parsed into infiltrated water, controlled by soil parameters, and surface runoff excess. Infiltrated precipitation drives subsurface water storage, which is utilized by trees, and simultaneously drives the baseflow component of total streamflow (Figure 6). Theoretically, the stormflow component should be composed only of overland

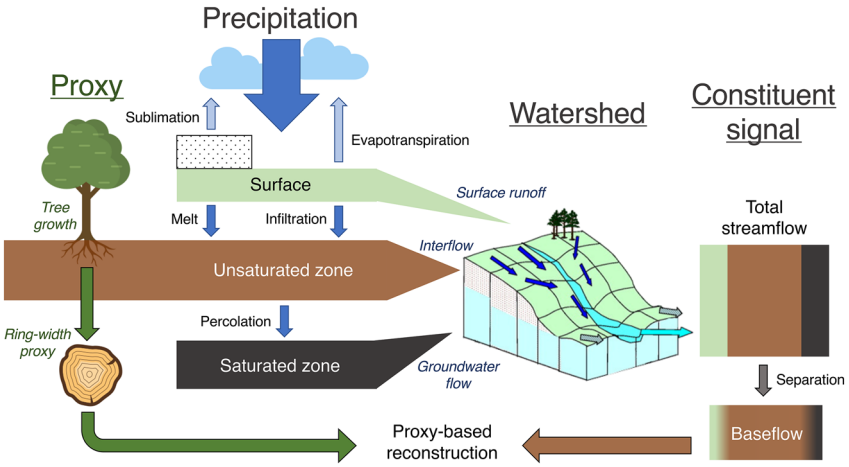

Figure 6. Proxy-based flow reconstruction approach superimposed on the hydrological cycle. Processes are separated by domain or scale, and color.

runoff, which does not interact with subsurface water storage used by trees and should have no causal link to tree growth. This matches our finding of few significant chronology correlations with the stormflow component (Figure 1). Rather than a causal link, the correlation peaks are likely due to imperfect baseflow separation and the spurious correlation caused by precipitation acting as a confounding variable for both baseflow and stormflow. This causal framework also helps to explain why the residual time series of "independent" stormflow from the reconstructed and instrumental data are not correlated. Once the influence of baseflow is removed, stormflow has no remaining information, further confirming that this is correlation without causation. This is the result one would expect from the unidirectional causality proposed in Figure 6, where baseflow and stormflow share a common driver, but where baseflow cannot directly influence stormflow. Because base and streamflow are highly correlated (r = 0.937/0.867 for DJF and JJA, respectively, 1931–1980), the difference in counts of significant correlations is not extreme (Table 1). However, if the relationship between tree growth and stormflow is random (beyond the correlation between base and stormflow), directly reconstructing total streamflow (baseflow + stormflow) risks adding stormflow "noise" to the baseflow "signal," thus degrading the signal-to-noise ratio.

As described above, baseflow and stormflow are driven by different hydrological processes and thus operate on different timescales. We suggest that tree-ring proxies in the Mid-Atlantic are tuned strongly to one of these constituent flow signals, e.g., baseflow, rendering the other constituent (stormflow) as correlated "noise" in a full streamflow reconstruction, related by the confounding variable of precipitation (Cook & Kairiukstis, 1990). This stormflow noise can represent a large proportion of total streamflow and have a highly skewed distribution because once soil is saturated, an additional unit of precipitation will have a negligible effect on subsurface water. By calibrating on a target with a greater amount of noise, there is a risk that the noise can mask the target signal and impart more information on the precalibration estimates, leading to weaker verification statistics and greater uncertainties associated with the reconstruction (Wigley et al., 1984). Simply put, the baseflow model (M3) calibrates on a less noisy target hydrologically closer to the tree-ring proxy and adds random noise to the reconstruction postcalibration, mimicking the assumed physical processes (Figure 6), while the streamflow models (M1 and M2) calibrate on a target that includes the full amount of noise. The lack of stormflow signal in the selected tree-ring chronologies is supported by residual analysis (Figure 5).

Ultimately, the degradation of signal could decrease correlation enough to prevent chronologies from being selected in the initial screening process and therefore a baseflow reconstruction with noise added postcalibration may increase the predictor pool and perform better than using streamflow as a predictand. As shown with the Potomac River flow for two different seasons, the number of chronologies that are significantly correlated with baseflow is higher than for streamflow (Table 1). The potential bias of performing PCA on a smaller subset of time series is evident when comparing the resulting reconstructions from a "discrete" seasonal approach (chronologies may only be used for one of the two seasons) to one using all significant

chronologies (including those with higher correlation with the alternate season flow). The reconstruction of JJA explains less variance than that of DJF (Table 2), despite covering part of the growing season of the Mid-Atlantic region. The lower statistics may be explained, in part, by the selection threshold applied. As with DJF, the relationship between tree growth and baseflow is stronger than with tree-ring growth and total streamflow. Correlations with early period USGS data (1901–1930) are higher for JJA compared to the DJF reconstructions; however, correlations between instrumental and reconstructed streamflow are significantly higher when targeting baseflow (M3), than when targeting total streamflow (M1/M2; Table 2). For M3, the precalibration and calibration correlations suggest a relatively stable relationship between tree-ring variability and hydroclimate.

Although the discrete DJF and JJA reconstructions display higher correlation than that of the instrumental data, it is lower than for reconstructions that allow overlapping predictors. The significant difference in correlation indicates that taking a "discrete" approach targets the independent variability to a greater extent and that any added correlation for the "all" approach allowing seasonally shared predictors stems mainly, or solely, from the shared predictor pool. The resulting JJA reconstructions from Models 1 and 2, which contain only 10 predictor chronologies, display significant trend over the past 250 years (Figure S4). The streamflow reconstruction produced from M3, with a higher number of predictors, does not contain any significant trend for the same period. When allowing chronologies to be predictors for both seasons, the total number of chronologies for M1/M2 increases to 17 and the resulting reconstructions no longer contain any significant positive trend. In addition to the higher number of chronologies going into the model, there may be further value in having separate baseflow reconstructions (e.g., baseflow is thought to have a stronger connection to the chemical water cycle (Koskelo et al., 2012). It is also possible that future work could produce predictors tuned to stormflow, and even at less fine resolutions these could be studied in tandem (e.g., Toomey et al., 2019).

4.3. Seasonal Reconstructions of Potomac River Streamflow

The DJF reconstruction of Potomac River flow is the first of its nature. Tree-ring based reconstructions of eastern North American winter hydroclimate are relatively rare, largely due to local climatology and the biophysical response of the trees (e.g., Cook & Jacoby, 1977). Unlike the western United States and Mexico, where growing-season soil moisture is reliant on winter and early spring precipitation (e.g., Stahle et al., 2009), rainfall during late spring and the summer is the main hydroclimatic driver of tree growth east of the hundredth meridian, 100°W (D'Orangeville et al., 2018; St. George, 2014) However, we argue that there are biophysical and hydroclimatic reasons for why tree rings in the Mid-Atlantic region displays a significant relationship with winter flow. Persistence between streamflow and soil moisture is evident, with Potomac River winter streamflow significantly correlated with drought conditions during the early growing season over a large area of the basin (as measured by May PDSI; Figure S1a), thus providing a feasible process link between tree growth and winter streamflow. The region of positive correlation broadly matches the spatial extent of significant DJF tree-ring chronologies, extending further than the relationship between summer streamflow and summer PDSI (Figure S1b), most likely due to the more localized convective nature of summer rainfall. Potomac River DJF baseflow can explain over 40% of the variance in May PDSI for several grid points in the catchment area. The relationship between streamflow and soil moisture is further strengthened by general temporal autocorrelation in both variables, although the magnitude varies between climatologies. For the Potomac River, stream and baseflow in February are significantly correlated with flow in June ($r = 0.388$ and 0.421, respectively, during the calibration period).

The persistence in the flow-soil moisture relationship is also reflected in correlations between baseflow and tree-ring chronologies from the Mid-Atlantic. The average signal is modest but the common response in tree-ring chronologies that are significantly correlated with Potomac River baseflow variability allows for past estimates of flow. All three models, including M3 which is only calibrated to baseflow, explain over 35% of the variance in DJF streamflow variability for the calibration period (Table 2). Comparisons with the early, precalibration USGS data reveal weak albeit positive correlations. Extreme years prior to the calibration period match historical records of flood and drought. Among these years is 1839, the lowest year of reconstructed DJF flow, for which eastern United States drought conditions have been described during the infamous "Trail of Tears" (Perdue & Green, 2007). Early observations from the Washington, DC, GHCN

precipitation gauge USW00093725 (Menne et al., 2012), highlight 1882 and 1884 as heavy precipitation winters—years that both rank in the 90th percentile of the DJF reconstruction. The second highest reconstructed flow is recorded for 1903, during which both the Potomac and Shenandoah rivers experienced March floods according to early USGS gages. Despite lower interannual correlation with early verification data, these estimates represent the first reconstruction of winter flow of the Potomac River and capture information about past extreme events. During the instrumental period (1930–2018), there is one episode of six consecutive years below-mean flow (during the 1960s drought). The DJF reconstruction records four such episodes during the 19th century: 1870–1875; 1852–1858; 1835–1841; 1818–1823 (Figures 4a and S2a). These periods suggest that the observations of Potomac River streamflow from the 20th and early 21st century may not encompass the full natural variability of the system.

Perhaps the most notable feature of the JJA reconstruction is the extended record of below-mean flows starting around 1840 and continuing until the mid-1870s (Figures 4b and S2b). Although individual years reach above the long-term mean, the average baseflow from 1837 to 1874 falls below the 20th percentile of the instrumental data. The 1960s drought appears to be the worst consecutive 5+ year period in the reconstruction; however, a similar extended period of summer droughts appears to have occurred in the 1820s. We suggest that these results should be considered by regional water resources management as they can provide baseline thresholds for future drought scenarios. Similarly, the 1830s high flows suggested by M3 appear to have been greater than any period reconstructed during the 20th century.

Streamflow for DJF and JJA is significantly correlated in the instrumental data but this natural seasonal correlation is slightly exaggerated in the reconstructions. The increased correlation is likely to stem from the biophysical persistence in the tree-ring chronologies; however, the reconstructions still offer insights on dual season low and high-flow. Several examples of accurate seasonality are present in the reconstruction. The JJA estimates record extreme low flows (<5th percentile) for 1819. The conditions do not appear to have improved over the subsequent winter (DJF flow for 1820 < 10th percentile). In one of many letters penned by Thomas Jefferson on the conditions, on May 15, 1820, he wrote (Looney, 2019):

> "... [the banks'] fatal effect has been greatly aggravated in this state by an unexampled drought, which having *prevailed* from June last to this time, destroyed the bread of that year, & threatens that of the present."

The reconstruction captures this paired JJA drought continuing into the winter season. A historical example of the opposite conditions is captured by the reconstructions for 1889. The reconstructions indicate high flows for both DJF and JJA (>90th percentile), resulting in flood conditions that caused over $1m worth of destruction to property on the C&O Canal between Cumberland, MD, and Washington, DC. This year also stands out in other proxy records that record flood-level flows (Toomey et al., 2019; Yanosky, 1983).

4.4. Limitations and Future Work

The work presented in this paper should not be considered replacements of existing reconstructions, but rather serve to demonstrate the value of baseflow separation in paleoclimatic reconstruction exercises. This is particularly important when proxy records are more closely attuned to one of these components—baseflow capturing near-surface groundwater or stormflow capturing precipitation extremes or individual events. Previous reconstructions have used different seasonal targets, a greater number of predictors (i.e., not limited to PC1 for year t), and different approaches to model the regression. Maxwell et al. (2011; 2017) produced significantly higher calibration statistics for the most replicated nest; however, it is worth noting that M3 presented here displays slightly higher correlation with precalibration USGS data (1901–1930; Maxwell et al., 2011/2017 $r = 0.441/0.401$ versus M3 $r = 0.551$). The similar results are produced despite M3 using a smaller pool of tree-ring chronologies, fewer predictors (i.e., only PC1), and fitting solely baseflow with an uncalibrated climatological portion added for the stormflow component. The Maxwell et al. (2011, 2017) reconstructions and the M3 warm season reconstruction are significantly correlated for the calibration period (1931–1980; $r = 0.627/0.772$), less than instrumental JJA and MJJAS streamflow ($r = 0.823$) but in line with what could be expected from the explained variances of the two reconstructions. The relationship is somewhat weaker for the strongest nest of the preinstrumental period (1750–1930; $r = 0.426/0.539$).

A major limitation of the reconstructions presented here, and of dendroclimatic reconstructions in general, is a short calibration period. Many of the chronologies selected for PCA in this analysis, also used in the Maxwell et al. (2011, 2017) reconstructions, were produced in the early 1980s (e.g., Cook & Jacoby, 1983), which is why the end year of calibration is 1980. Extending these high-quality chronologies could provide more robust models. Producing new chronologies from multiple species (Maxwell et al., 2015) and the development of EW and LW series from existing/new chronologies could provide increased skill, especially for the seasonal resolution of past estimates. Additionally, novel approaches to address potential sign biases in streamflow reconstructions have recently been developed (Robeson et al., 2020). Such methods, as well as further analyses on the nonstability of climate signals in tree rings, could further improve estimates of past flow.

Future work elsewhere may focus on stormflow in addition to, or rather than, baseflow depending on the local climatology. Records of tree growth have been used to reconstruct short-term precipitation events (e.g., Howard & Stahle, 2020) and such records most likely contain a strong stormflow signal. The spatial correlation of flow in snowpack driven basins has been shown to have significant skill in predicting downstream gages (Ravindranath et al., 2019) and similarly, independent seasonal reconstructions may add skill from temporal persistence. It is likely that there are locations where both baseflow and stormflow could be reconstructed separately if utilizing both EW and LW records of multiple species. By doing so, it may be possible to minimize the true uncertainty associated with paleoclimatic reconstruction of the total streamflow.

5. Conclusions

Tree-ring chronologies from the Mid-Atlantic region consistently display higher correlations with baseflow than with total streamflow variability of the Potomac River for both cool (DJF) and warm (JJA) periods. Few chronologies are correlated with stormflow despite high correlation between baseflow and stormflow, supporting the hypothesis that baseflow is more closely tied to tree-ring proxies in this region and that the stormflow component is largely noise, from a dendroclimatological perspective. Models calibrated on baseflow produce stronger reconstruction statistics than for streamflow, presumably for the same reason. For the warm season, the reconstruction skill prior to the calibration period is significantly stronger when using baseflow as the initial target. Combined with alternative approaches to the regression process (e.g., a GLM gamma-linked regression), we suggest that calibrating on one of the theoretical constituents of flow (in this case baseflow) can help improve the resulting reconstruction and provide more accurate uncertainty estimates. The new reconstructions presented here represent conservative estimates of both winter and summer Potomac River streamflow, using only a single predictor (PC1) and enforcing a discrete predictor rule that no chronologies are shared between winter and summer reconstructions. Qualitative comparisons with historical records prior to the calibration period indicate that both reconstructions contain information on past extreme flows. Even in the most recent 200 years, for which the reconstructions have the greatest replication, flows outside the magnitude range of the calibration period are recorded. We hypothesize that the separation of flow into the theoretical constituents of base and storm components can provide useful information in fundamentally different hydrological settings, such as the western United States.

Data Availability Statement

The seasonal reconstructions of baseflow and streamflow are available through the NOAA NCDC Paleoclimatology data bank (https://www.ncdc.noaa.gov/data-access/paleoclimatology-data).

Acknowledgements
The authors thank Cherie Schultz for discussions prior to and during the writing of this manuscript, as well as three anonymous reviewers for constructive comments. The authors appreciate all researchers who have contributed, and continue to contribute, to the ITRDB.

References

Ahmed, S. N., Bencala, K. R., & Schultz, C. L. (2015). *2015 Washington metropolitan area water supply study—Demand and resource availability forecast for the year 2040 (ICPRB Rep. 15-4)*. Rockville, MD: Interstate Commission on the Potomac River Basin.

Aloni, R. (1991). Wood formation in deciduous hardwood trees. In A. S. Raghavendera (Ed.), *Physiology of trees* (pp. 175–197). New York, NY: John Wiley & Sons.

Bhatkoti, R., Triantis, K., Moglen, G. E., & Sabounchi, N. S. (2018). Performance assessment of a water supply system under the impact of climate change and droughts: Case study of the Washington Metropolitan Area. *Journal of Infrastructure Systems*, *24*, 05018002. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000435

Cleaveland, M. K., & Stahle, D. W. (1989). Tree-ring analysis of surplus and deficit runoff in the White River, Arkansas. *Water Resources Research*, *25*, 1391–1401. https://doi.org/10.1029/WR025i006p01391

Cook, E. R., & Jacoby, G. C. (1977). Tree-ring-drought relationship in the Hudson Valley, New York. *Science*, *198*, 399–401. https://doi.org/10.1126/science.198.4315.399
Cook, E. R., & Jacoby, G. C. (1983). Potomac River streamflow since 1730 as reconstructed by tree rings. *Journal of Climate and Applied Meteorology*, *22*, 1659–1672. https://doi.org/10.1175/1520-0450(1983)022<1659:PRSSAR>2.0.CO;2
Cook, E. R., & Kairiukstis, L. A. (1990). *Methods of dendrochronology: Applications in the environmental sciences*. (pp. 394). Dordrecht:Kluwer Academic Publishers.
Cook, E. R., Meko, D. M., Stahle, D. W., & Cleaveland, M. K. (1999). Drought reconstructions of the continental United States. *Journal of Climate 12*, 1145–1162. https://doi.org/10.1175/1520-0442(1999)012<1145:DRFTCU>2.0.CO;2
D'Orangeville, L., Maxwell, J., Kneeshaw, D., Pederson, N., Duchesne, L., Logan, T., et al. (2018). Drought timing and local climate determine the sensitivity of eastern temperate forests to drought. *Global Change Biology*, *24*, 2339–2351. https://doi.org/10.1111/gcb.14096
Devineni, N., Lall, U., Pederson, N., & Cook, E. R. (2013). A tree-ring based reconstruction of Delaware River basin streamflow using hierarchical Bayesian regression. *Journal of Climate*, *26*, 4357–4374. https://doi.org/10.1175/JCLI-D-11-00675.1
Eckhardt, K. (2008). A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. *Journal of Hydrology*, *352*, 168–173. https://doi.org/10.1016/j.jhydrol.2008.01.005
Enfield, D. B., Mestas-Nuñez, A. M., & Trimble, P. J. (2001). The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. *Geophysical Research Letters*, *28*, 2077–2080. https://doi.org/10.1029/2000GL012745
Fisher, R. A. (1921). On the "probable error" of a coefficient of correlation deduced from a small sample. *Metron*, *1*, 3–32.
Gonzales, A. L., Nonner, J., Heijkers, J., & Uhlenbrook, S. (2009). Comparison of different base flow separation methods in a lowland catchment. *Hydrology and Earth System Sciences*, *13*, 2055–2068. https://doi.org/10.5194/hess-13-2055-2009
Gonzales, J., & Valdes, J. (2003). Bivariate drought recurrence analysis using tree ring reconstructions. *Journal of Hydrologic Engineering*, *8*, 247–258. https://doi.org/10.1061/(ASCE)1084-0699(2003)8:5(247)
Gou, X., Chen, F., Cook, E., Jacoby, G., Yang, M., & Li, J. (2007). Streamflow variations of the Yellow River over the past 593 years in western China reconstructed from tree rings. *Water Resources Research*, *43*, W06434. https://doi.org/10.1029/2006WR005705
Hagen, E. R., Holmes, K. J., Kiang, J. E., & Steiner, R. C. (2005). Benefits of iterative water supply forecasting in the Washington, D.C. Metropolitan area. *Journal of the American Water Resources Association*, *41*, 1417–1430. https://doi.org/10.1111/j.1752-1688.2005.tb03809.x
Helsel, D. R., Hirsch, R. M., Ryberg, K. R., Archfield, S. A., & Gilroy, E. J. (2020). Statistical methods in water resources. In U.S. Geological Survey techniques and methods (Book Vol. 4, Chapter A3, pp. 456). https://doi.org/10.3133/tm4a3
Ho, M., Lall, U., & Cook, E. R. (2016). Can a paleodrought record be used to reconstruct streamflow?: A case study for the Missouri River Basin. *Water Resources Research*, *52*, 5195–5212. http://doi.org/10.1002/2015WR018444
Howard, I. M., & Stahle, D. W. (2020). Tree-ring reconstructions of single day precipitation totals over eastern Colorado. *Monthly Weather Review*, *148*, 597–612. https://doi.org/10.1175/MWR-D-19-0114.1
Jolliffe, I. T. (2002). Principal Components in Regression Analysis. *Principal Component Analysis*, Springer Series in Statistics2(167–198). New York, NY: Springer. https://doi.org/10.1007/b98835
Koskelo, A. I., Fisher, T. R., Utz, R. M., & Jordan, T. E. (2012). A new precipitation-based method of baseflow separation and event identification for small watersheds (<50km2). *Journal of Hydrology*, *450-451*, 267–278. https://doi.org/10.1016/j.jhydrol.2012.04.055
Koutsoyiannis, D., & Montanari, A. (2007). Statistical analysis of hydroclimatic time series: Uncertainty and insight. *Water Resources Research*, *43*, W05429. https://doi.org/10.1029/2006WR005592
Looney, J. J. (2019). *The papers of Thomas Jefferson: Retirement series*. (Vol. 15). Princeton, NJ: Princeton University Press.
Lorenz, D. (2017). *DVstats: Functions to manipulate daily-values data (R-package)*. Washington, DC: US Geological Survey.
Lott, D. A., & Stewart, M. T. (2016). Base flow separation: A comparison of analytical and mass balance methods. *Journal of Hydrology*, *535*, 525–533. https://doi.org/10.1016/j.jhydrol.2016.01.063
Maxwell, J. T., Harley, G. L., & Matheus, T. J. (2015). Dendroclimatic reconstructions from multiple co-occurring species: A case study from an old-growth deciduous forest in Indiana, USA. *International Journal of Climatology*, *15*, 860–870. https://doi.org/10.1002/joc.4021
Maxwell, R. S., Harley, G. L., Maxwell, J. T., Rayback, S. A., Pederson, N., Cook, E. R., et al. (2017). An interbasin comparison of tree-ring reconstructed streamflow in the eastern United States. *Hydrological Processes*, *31*, 2381–2394. https://doi.org/10.1002/hyp.11188
Maxwell, R. S., Hessl, A. E., Cook, E. R., & Pederson, N. (2011). A multispecies tree ring reconstruction of Potomac River streamflow (950–2001). *Water Resources Research*, *47*, W05512. https://doi.org/10.1029/2010WR010019
Meko, D. M., Friedman, J. M., Touchan, R., Edmondson, J. R., Griffin, E. R., & Scott, J. A. (2015). Alternative standardization approaches to improving streamflow reconstructions with ring-width indices of riparian trees. *The Holocene*, *25*, 1093–1101. https://doi.org/10.1177/0959683615580181
Meko, D. M., Panyushkina, I. P., Agafanov, L., & Edwards, J. A. (2020). Impact of high flows of an Arctic river on ring widths of floodplain trees. *The Holocene*. https://doi.org/10.1177/0959683620902217
Meko, D., Therrell, M., Baisan, C., & Hughes, M. (2001). Sacramento River flow reconstructed to A.D. 869 from tree rings. *Journal of the American Water Resources Association*, *37*, 1029–1039. https://doi.org/10.1111/j.1752-1688.2001.tb05530.x
Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology*, *29*(7), 897–910.
Palmer, W. C. (1965). *Meteorological drought. Weather Bureau research paper 45*. Washington, DC: US Department of Commerce.
Partington, D., Brunner, P., Simmons, C. T., Werner, A. D., Therrien, R., Maier, H. R., & Dandy, G. C. (2012). Evaluation of outputs from automated baseflow separation methods against simulated baseflow from a physically based, surface water-groundwater flow model. *Journal of Hydrology*, *458–459*, 28–39. https://doi.org/10.1016/j.jhydrol.2012.06.029
Perdue, T., & Green, M. D. (2007). *The Cherokee Nation and the trail of Tears* (pp. 189). Penguin Publishing.
Pettyjohn, W. A., & Henning, R. (1979). *Preliminary estimate of ground-water recharge rates, related streamflow and water quality in Ohio*. Washington, DC: U.S. Department of the Interior.
Piggott, A. R., Moin, S., & Southam, C. (2005). A revised approach to the UKIH method for the calculation of baseflow. *Hydrological Sciences Journal*, *50*, 911–920. https://doi.org/10.1623/hysj.2005.50.5.911
Prairie, J., Rajagopalan, B., Fulp, T., & Zagona, E. (2006). Modified K-NN model for stochastic streamflow simulation. *Journal of Hydrologic Engineering*, *11*, 371–378. http://doi.org/10.1061/(ASCE)1084-0699(2006)11:4(371)
Rao, M. P., Cook, E. R., Cook, B. I., Palmer, J. G., Uriarte, M., Devineni, N., et al. (2018). Six centuries of Upper Indus basin streamflow variability and its climatic drivers. *Water Resources Research*, *54*, 5687–5701. https://doi.org/10.1029/2018WR023080
Ravindranath, A., Devineni, N., Lall, U., Cook, E. R., Pederson, G., Martin, J., & Woodhouse, C. (2019). Streamflow reconstruction in the Upper Missouri River basin using a novel Bayesian network model. *Water Resources Research*, *55*, 7694–7716. https://doi.org/10.1029/2019WR024901

Razavi, S., Elshorbagy, A., Wheater, H., & Sauchyn, D. (2015). Towards understanding nonstationarity in climate and hydrology through tree ring proxy records. *Water Resources Research*, *51*, 1813–1830. https://doi.org/10.1002/2014WR015696

Rice, K. C., & Hirsch, R. M. (2012). *Spatial and temporal trends in runoff at long-term streamgages within and near the Cheasepeake Bay watershed*. Reston, VA: US Geological Survey.

Rice, J. L., Woodhouse, C., & Lukas, J. (2009). Science and decision-making: Water management and tree-ring data in the western United States. *Journal of the American Water Resources Association*, *45*, 1248–1259. https://doi.org/10.1111/j.1752-1688.2009.00358.x

Robeson, S. M., Maxwell, J. T., & Ficklin, D. L. (2020). Bias correction of paleoclimatic reconstruction: A new look at 1200 years of upper Colorado River flow. *Geophysical Research Letters*, *47*, e2019GL086689. https://doi.org/10.1029/2019GL086689

Ruel, J. J., & Ayres, M. P. (1999). Jensen's inequality predicts effects of environmental variation. *Trends in Ecology and Evolution*, *14*, 361–366. https://doi.org/10.1016/S0169-5347(99)01664-X

Saito, L., Biondi, F., Salas, J. D., Panorska, A. K., & Kozubowski, T. J. (2008). A watershed modeling approach to streamflow reconstruction from tree-ring records. *Environmental Research Letters*, *3*, 024006. https://doi.org/10.1088/1748-0326/3/2/024006

Sauchyn, D., & Ilich, N. (2017). Nine hundred years of weekly streamflows: Stochastic downscaling of ensemble tree-ring reconstructions. *Water Resources Research*, *53*, 9266–9283. https://doi.org/10.1002/2017WR021585

Schultz, C. L., Ahmed, S. N., Mandel, R., & Moltz, H. L. N. (2014). Improvements in HSPF's low-flow predictions by implementation of a power law groundwater storage-discharge relationship. *Journal of the American Water Resources Association*, *50*, 909–927. https://doi.org/10.1111/jawr.12144

Schultz, C. L., Ahmed, S. N., Moltz, H. L. N., & Seck, A. (2017). *Washington Metropolitan area water supply alternatives—Meeting the challenge of growth and climate Change (ICPRB Rep. 17-3)*. Rockville, MD: Interstate Commission on the Potomac River Basin.

Sloto, R. A., & Crouse, M. Y. (1996). HYSEP: A computer program for streamflow hydrograph separation and analysis. *USGS water resources investigations (Rep. 96-4040*, pp. 46).

Stagge, J. H., & Moglen, G. E. (2013). A nonparametric stochastic method for generating daily climate-adjusted streamflows. *Water Resources Research*, *49*, 6179–6193. https://doi.org/10.1002/wrcr.20448

Stagge, J. H., & Moglen, G. E. (2017). Water resources adaptation to climate and demand change in the Potomac River. *Journal of Hydrologic Engineering*, *22*, 04017050. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001579

Stagge, J. H., Rosenberg, D. E., DeRose, R. J., & Rittenour, T. M. (2018). Monthly paleostreamflow reconstruction from annual tree-ring chronologies. *Journal of Hydrology*, *557*, 791–804. https://doi.org/10.1016/j.jhydrol.2017.12.057

Stahle, D. W., Cleaveland, M. K., Grissino-Mayer, H., Griffin, R. D., Fye, F. K., Therrell, M. D., et al. (2009). Cool and warm season precipitation reconstructions over western New Mexico. *Journal of Climate*, *22*, 3729–3750. https://doi.org/10.1175/2008JCLI2752.1

Stahle, D. W., Cook, E. R., Burnette, D. J., Torbenson, M. C. A., Howard, I. M., Griffin, D., et al. (2020). Dynamics, variability, and change in seasonal precipitation reconstructions for North America. *Journal of Climate*, *33*, 3173–3195. https://doi.org/10.1175/JCLI-D-19-0270.1

St. George, S. (2014). An overview of tree-ring width records across the Northern Hemisphere. *Quaternary Science Reviews*, *95*, 132–150. https://doi.org/10.1016/j.quascirev.2014.04.029

St. George, S., Meko, D. M., & Cook, E. R. (2010). The seasonality of precipitation signals embedded within the North American Drought Atlas. *The Holocene*, *20*, 983–988. https://doi.org/10.1177/0959683610365937

Strange, B. M., Maxwell, J. T., Robeson, S. M., Harley, G. L., Therrell, M. D., & Ficklin, D. L. (2019). Comparing three approaches to reconstructing streamflow using tree rings in the Wabash River basin in the Midwestern, US. *Journal of Hydrology*, *573*, 829–840. https://doi.org/10.1016/j.jhydrol.2019.03.057

Tebaldi, C., Hayhoe, K., Arblaster, J. M., & Meehl, G. A. (2006). Going to the extremes. *Climatic Change*, *79*, 185–211. https://doi.org/10.1007/s10584-006-9051-4

Tingstad, A. H., Groves, D. G., & Lempert, R. J. (2014). Paleoclimate scenarios to inform decision making in the water resource management: Example from southern California's Inland Empire. *Journal of Water Resources Planning and Management*, *140*. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000403

Toomey, M., Cantwell, M., Colman, S., Cronin, T., Donnelly, J., Giosan, L., et al. (2019). The Mighty Susquehanna—Extreme floods in eastern North America during the past two millennia. *Geophysical Research Letters*, *46*, 3398–3407. https://doi.org/10.1029/2018GL080890

Torbenson, M. C. A., & Stahle, D. W. (2018). The relationship between cool and warm season moisture over the central United States, 1685–2015. *Journal of Climate*, *31*, 7909–7924. https://doi.org/10.1175/JCLI-D-17-0593.1

Urrutia, R. B., Lara, A., Villalba, R., Christie, D. A., Le Quesne, C., & Cuq, A. (2011). Multicentury tree ring reconstruction of annual streamflow for the Maule River watershed in south central Chile. *Water Resources Research*, *47*, W06527. https://doi.org/10.1029/2010WR009562

Wigley, T. M., Briffa, K. R., & Jones, P. D. (1984). On the average value of correlated time series, with applications in dendroclimatology and hydrometeorology. *Journal of Climate and Applied Meteorology*, *23*, 201–213. https://doi.org/10.1175/1520-0450(1984)023<0201:OTAVOC>2.0/CO;2

Woodhouse, C. A., Gray, S. T., & Meko, D. M. (2006). Updated streamflow reconstructions for the Upper Colorado River basin. *Water Resources Research*, *42*, W05415. https://doi.org/10.1029/2005WR004455

Woodhouse, C. A., Lukas, J., Morino, K., Meko, D., & Hirschboeck, K. (2016). *Using the past to plan for the future? The value of paleoclimate reconstructions for water resource planning.* In Water policy and planning in a variable and changing climate (pp. 161–182). Indianapolis, IN: Taylor & Francis.

Woodhouse, C. A., Meko, D. M., Griffin, D., & Castro, C. L. (2013). Tree rings and multiseason drought variability in the lower Rio Grande Basin, USA. *Water Resources Research*, *49,* 844–850. https://doi.org/10.1002/wrcr.20098

Woodhouse, C. A., & Pederson, G. T. (2018). Investigating runoff efficiency in Upper Colorado River streamflow over past centuries. *Water Resources Research*, *54,* 286–300. https://doi.org/10.1002/2017WR021663

Yang, Q., Tian, H., Friedrichs, M. A. M., Liu, M., Li, X., & Yang, J. (2015). Hydrological responses to climate and land-use changes along the north American east coast: A 110-year historical reconstruction. *Journal of American Water Resources Association*, *51*, 47–67. https://doi.org/10.1111/jawr.12232

Yanosky, T. M. (1983). Evidence of floods on the Potomac River from anatomical abnormalities in the wood of flood-plain trees (US Geol. Surv. Prof. Pap. 1296). https://doi.org/10.3133/pp1296

Zhao, S., Pederson, N., D'Orangeville, L., HilleRisLambers, J., Boose, E., Penone, C., et al. (2019). The International Tree-Ring Data Bank (ITRDB) revisited: Data availability and global ecological representativity. *Journal of Biogeography*, *46*, 355–368. https://doi.org/10.1111/jbi.13488