Check for updates

AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2020WR029394

Multi-Proxy, Multi-Season Streamflow Reconstruction With Mass Balance Adjustment

Hung T. T. Nguyen[1,2], **Stefano Galelli**[1,2], **Chenxi Xu**[3,4], **and Brendan M. Buckley**[2]

[1]Pillar of Engineering Systems and Design, Singapore University of Technology and Design, Singapore, Singapore, [2]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA, [3]Key Laboratory of Cenozoic Geology and Environment, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China, [4]CAS Center for Excellence in Life and Paleoenvironment, Beijing, China

Key Points:

- Ring width and $\delta^{18}O$ are combined to reconstruct dry season, wet season, and annual streamflow
- Optimal proxy combinations are found with an automatic input selection scheme
- Mass balance adjustment improves the agreement between seasonal and annual reconstructions

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

H. T. T. Nguyen,
hnguyen@ldeo.columbia.edu

Citation:

Nguyen, H. T. T., Galelli, S., Xu, C., & Buckley, B. M. (2021). Multi-proxy, multi-season streamflow reconstruction with mass balance adjustment. *Water Resources Research*, *57*, e2020WR029394. https://doi.org/10.1029/2020WR029394

Received 4 DEC 2020
Accepted 16 JUL 2021

© 2021. American Geophysical Union. All Rights Reserved.

Abstract Despite having offered important hydroclimatic insights, streamflow reconstructions still see limited use in water resources operations, because annual reconstructions are not suitable for decisions at finer time scales. The few attempts toward sub-annual reconstructions have relied on statistical disaggregation, which uses none or little proxy information. Here, we develop a novel framework that optimizes proxy combinations to simultaneously produce seasonal and annual reconstructions. Importantly, the framework ensures that total seasonal flow matches annual flow closely. This mass balance criterion is necessary to avoid misguiding water management decisions, such as the allocation of water rights or dam release decisions. Using the framework, and leveraging a multi-species network of ring width and cellulose $\delta^{18}O$ in Southeast Asia, we reconstruct seasonal and annual inflow to Thailand's largest reservoir. The reconstructions are statistically skillful. Furthermore, they preserve the mass balance well: the differences are mostly within 10% of the mean annual flow. As a result, the reconstructions provide more reliable estimates of the seasonal and annual surface water availability. This work is one step closer toward operational usability of streamflow reconstruction in water resources management.

Plain Language Summary Long history of river discharge, or streamflow, can be reconstructed from tree rings. These reconstructions help us understand the water cycle in the past, but they have not been widely used in water resources operations. This is because reconstructions are often annual (having only one data point per year). By combining different tree species and different features of tree rings (for example, ring width and stable isotope ratio), it is possible to reconstruct seasonal streamflow in addition to the annual one, and that is our first goal. But a major challenge arises: how do we ensure that the total flow volume of the seasonal reconstructions closely matches the annual one? This mass balance criterion is important to avoid misguiding water management decisions such as allocating water to different sectors. We develop a novel method to reconstruct seasonal and annual streamflow while accounting for mass balance at the same time. Our work is thus a step closer toward operational usability of streamflow reconstructions in water resources management.

1. Introduction

Dendrohydrology, the study of past hydroclimate using tree rings, has been largely motivated by water resources management. The field traces back to Hardman and Reil (1936), who recognized that instrumental records were too short to understand drought trends and demonstrated that better understanding could be gained from exploring the links between tree rings and streamflow. Their work was motivated by contemporary droughts in California that affected irrigation. At the same time, Hawley (1937) found strong correlations between tree rings and streamflow in Tennessee, USA, and was probably the first to show the lagged relationship of streamflow to tree rings. Also to understand droughts, Schulman (1945) established a tree ring chronology for the Colorado River Basin, this time motivated by the war effort—to examine Hoover Dam's hydropower production reliability to meet wartime demand. While these early works stopped at studying tree ring indices, dendrohydrology took a big step when Stockton (1971), leveraging advanced multivariate techniques (Fritts et al., 1971), showed that reconstructing streamflow record back in time was feasible—long term surface water availability could now be quantified directly. Soon, multiple streamflow reconstructions were developed across the Colorado River Basin (Stockton & Jacoby, 1976), revealing the

shortcomings of the Colorado River Compacts (Woodhouse et al., 2006), and providing insights about long term hydrology of Lake Powell, the United States' second largest reservoir.

Streamflow reconstruction has become "an important planning and research tool" in water resources management (Meko & Woodhouse, 2011). Yet, its use in practical, operational aspects of water management is still limited in scope and effectiveness (Galelli et al., 2021). That is because reconstructions often target specific components of the hydrograph that best correlate with tree ring proxies. Perhaps most commonly, reconstructions from ring width target the growth season (e.g., D'Arrigo, Abram, et al., 2011; Güner et al., 2017). Another example is given by reconstructions targeting peak flow using stable oxygen isotope ratio (δ^{18}O) of tree ring cellulose (C. Xu et al., 2019). These reconstructions reveal important insights about the hydroclimate, but do not provide the total annual surface water availability. Other works target the annual flow (e.g., Nguyen & Galelli, 2018; Rao et al., 2018), but even so, the annual resolution is not suitable for making operational decisions at finer time scales—crop planning, for instance, is often based on seasonal flow; reservoir releases are determined at monthly or even daily time steps.

The water resources community recognizes the need for sub-annual reconstructions. Attempts toward this goal have relied on statistical disaggregation, assuming some statistical relationships between the sub-annual and annual flows (Prairie et al., 2007, 2008; Sauchyn & Ilich, 2017). These assumptions are reasonable but not always valid (Figure S1). More importantly, paleoclimatic proxies are not used in these methods, and their rich information are not utilized. Recent progress was made by Stagge et al. (2018), who used multi-species chronologies as additional inputs to disaggregation, showing that these inputs can be weighted differently for each month to improve the monthly reconstructions.

The works of Stagge et al. (2018), C. Xu et al. (2019), and others discussed above suggest that different proxies have different seasonal sensitivities. Instead of disaggregation, multiple proxies can be used to simultaneously reconstruct sub-annual (e.g., seasonal) and annual flows. This hypothesis has just been tested with promising results by Wise (2021), who produced skillful monthly, quarterly, and semi-annual reconstructions for multiple Californian watersheds. However, two challenges remain in sub-annual reconstructions. How to combine proxies optimally for different targets? And how to ensure that the seasonal flows add up to the annual flow, that is, how to account for mass balance? We develop a unified framework to address both challenges. Mass balance is accounted for by a term in the regression formulation that penalizes the differences between total seasonal flow and annual flow (Section 3.1), and proxy combination is optimized with an automatic input selection scheme (Section 3.2). We test the framework with a case study in the Chao Phraya River Basin, Thailand, pooling together a multi-species network of ring width and cellulose δ^{18}O chronologies from Southeast Asia (Section 2). This work is one step closer toward operational usability of streamflow reconstruction in water resources management.

2. Study Site and Data

2.1. The Chao Phraya River Basin and Streamflow Data

The Chao Phraya Basin covers roughly 30% of Thailand's area (~150,000 km^2), but is its most important economic region (Figures 1a and 1b). The basin is home to about 25 million people. In 2005, the basin employed more than three-fourths of the country's workforce and generated about two-thirds of its GDP (Thomas, 2005). Water is the basin's lifeblood. The basin has about 1.45 million ha of irrigated land (Divakar et al., 2011), and 1.3 GW of installed hydropower capacity (Electricity Generating Authority of Thailand, 2013). Another 2.5 GW of thermal power also depends on river flow for cooling (Chowdhury et al., 2021). Yet, water is not abundant: freshwater availability *per capita* stands at about 2,230 m^3/year (Divakar et al., 2011; World Bank, 2011), less than the national average (3,244 m^3/year) and only 39% of the world's average (5,732 m^3/year) (FAO, 2017). This fundamental resource is mainly controlled by Bhumibol and Sirikit Dams, which create two large impoundments with active capacities of 9.7 and 6.7 km^3. Reliable operations of these reservoirs require accurate assessment of inflow availability, on both inter- and intra-annual scales—the latter is particularly important to address concerns about seasonal water availability as demanded by different sectors (Thomas, 2005).

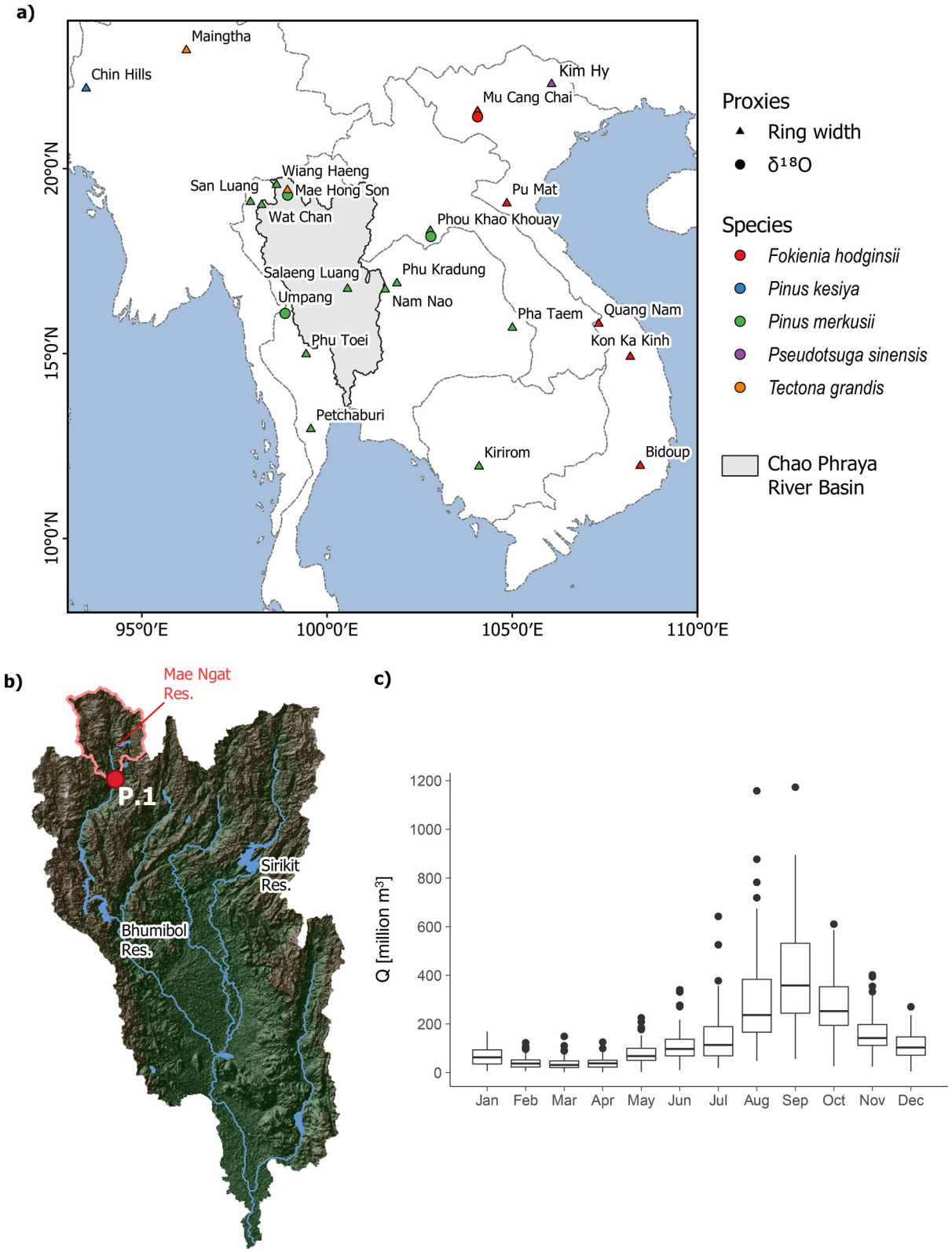

Figure 1. (a) Map of the study region, showing the Chao Phraya River Basin and the proxy network. (b) Topographic map of the Chao Phraya River Basin, the target streamflow station (P.1) with its drainage area, and reservoirs of interest. (c) Distribution of monthly streamflow at P.1.

The basin has a dominant monsoon climate, with most of the annual rainfall delivered by the Southwest Monsoon (early May–October–November). The seasonal and interannual patterns of water availability therefore mirror the dynamics of the Southeast Asian Monsoon, as well as its interaction with the El Niño–Southern Oscillation (ENSO). El Niño events tend shorten the rain season, while La Niñas bring more abundant precipitations (B. I. Cook & Buckley, 2009). It is indeed during, or shortly after, La Niña episodes that flood risk increases: when tropical storms during the later part of the monsoon season encounter an already wet catchment, we observe the ideal conditions for extreme floods, as in 1973 or 2005 (Lim & Boochabun, 2012).

The site of interest for seasonal and annual streamflow reconstructions is station P.1, located upstream of Bhumibol Reservoir (Figure 1b). Seasonal streamflow reconstructions at P.1 may potentially improve future assessments of the reservoir's inflow variability. P.1 is also a good candidate for reconstruction as it has the longest and most complete record in Thailand: daily data are available from April 1921 to present. The distribution of monthly streamflow at P.1 is shown in Figure 1c. Streamflow here is typical of this region: lowest flows often occur between February and April, while peak flows often occur in August to October. Flows in the peak season are highly varied, resulting from the complex interactions between monsoon rain and tropical storms that we described above.

Since 1985, the river upstream of P.1 has been impounded by the Mae Ngat Dam, which, at full capacity, stores about 14% of P.1's mean annual flow. Dam operations modify the seasonal streamflow patterns and may interfere with the proxy-streamflow relationship. Therefore, we naturalized the streamflow data from 1985. The naturalization process is described in Text S3. After naturalization, we aggregated daily data into dry season (November–June), wet season (July–October), and water year (November–October). The season delineation was determined by the method of B. I. Cook and Buckley (2009) (Text S2), and is consistent with a visual examination of the annual hydrograph (Figure 1c). To match the proxies' time span, we finally used the streamflow data from November 1921 to October 2005.

2.2. The Southeast Asian Dendrochronology Network

Over the past three decades, an extensive network of tree ring chronologies have been developed in Southeast Asia. This network has been instrumental in improving our understanding of Southeast Asia's hydroclimate and history. Tree ring data from Thailand and northern Vietnam (Buckley, Palakit, et al., 2007; Sano et al., 2009) revealed a multidecadal drought, what is later known as the Strange Parallel Droughts (E. R. Cook et al., 2010), which coincided with a tumultuous period of Southeast Asian history (Lieberman, 2003; Lieberman & Buckley, 2012). Further back in time, tree ring data from southern Vietnam linked megadroughts in the 14th and 15th centuries to the demise of the Angkor Civilization (Buckley et al., 2010, 2014). These findings show the intimate links between Southeast Asian societies and their water resources, and suggest the potential benefit of more hydroclimatic reconstructions in the region.

In this work, we use 20 tree ring chronologies from Vietnam, Laos, Cambodia, Thailand, and Myanmar (Figure 1). The chronologies at Kirirom, Petchaburi, Pha Taem, and Wiang Haeng are published here for the first time. The metadata of the chronologies are provided in Table 1. The common period of most chronologies in our network is 1748–2005 (Figure S5), and is the same as the time span of our δ^{18}O network. Several chronologies are some decades shorter. Following Stagge et al. (2018), we imputed the missing years using the R package missMDA (Josse & Husson, 2016) (see Figure S6). We imputed the tree ring data instead of building nested models because nesting is not applicable in our reconstruction framework. As we shall explain in Section 3.1, the framework is designed to account for mass balance, tuning the regression parameters such that the total sub-annual flow matches the annual flow closely. With nesting, the final variance correction can disrupt the mass balance.

2.3. Stable Oxygen Isotope Ratio (δ^{18}O) of Tree Ring Cellulose

We use four δ^{18}O chronologies that were developed in Laos, Thailand, and Vietnam over the past decade (Figure 1). Samples were collected and processed to alpha-cellulose, and δ^{18}O time series were measured in prior works (Table 2). Details of the laboratory procedure can be found in C. Xu et al. (2011). δ^{18}O exhibits

Table 1
Metadata of Tree Ring Width Chronologies

Site	Longitude	Latitude	Species	References
Bidoup	108.45	11.97	*Fokienia hodginsii*	Buckley et al. (2010)
Chin Hills	93.50	22.17	*Pinus kesiya*	Rao et al. (2020)
Kim Hy	106.04	22.25	*Pseudotsuga sinensis*	Hansen et al. (2017)
Kirirom	104.10	11.95	*Pinus merkusii*	This study[a]
Kon Ka Kinh	108.18	14.91	*Fokienia hodginsii*	Buckley et al. (2019)
Mae Hong Son	98.93	19.28	*Tectona grandis*	Buckley, Palakit, et al. (2007)
Maingtha	96.20	23.20	*Tectona grandis*	D'Arrigo, Palmer, et al. (2011)
Mu Cang Chai	104.06	21.40	*Fokienia hodginsii*	Sano et al. (2009)
Nam Nao	101.57	16.73	*Pinus merkusii*	Buckley et al. (1995)
Petchaburi	99.56	12.96	*Pinus merkusii*	This study
Pha Taem	105.00	15.70	*Pinus merkusii*	This study
Phou Khao Khouay	102.79	18.32	*Pinus merkusii*	Buckley, Duangsathaporn, et al. (2007)
Phu Kradung	101.88	16.90	*Pinus merkusii*	D'Arrigo et al. (1997)
Phu Toei	99.43	14.98	*Pinus merkusii*	E. R. Cook et al. (2010)
Pu Mat	104.85	19.06	*Fokienia hodginsii*	Buckley et al. (2019)
Quang Nam	107.33	15.81	*Fokienia hodginsii*	Buckley et al. (2017)
Salaeng Luang	100.55	16.75	*Pinus merkusii*	Buckley et al. (1995)
San Luang	97.93	19.10	*Pinus merkusii*	E. R. Cook et al. (2010)
Wat Chan	98.23	19.02	*Pinus merkusii*	Buckley et al. (1995)
Wiang Haeng	98.64	19.56	*Pinus merkusii*	This study

[a]Several cores from this site were analyzed by Zhu et al. (2012) for $\delta^{18}O$ but the ring width chronology has not been published until now.

strong mechanical and statistical relationship with the hydroclimate (Sano et al., 2012; C. Xu et al., 2011), and has been used to reconstruct wet season precipitation in the region (C. Xu et al., 2015, 2018). $\delta^{18}O$ in Laos was also found to have significant negative correlation with Mekong River's water level (C. Xu et al., 2013), suggesting promising hydrological applications. Finally, three $\delta^{18}O$ chronologies were used to reconstruct Chao Phraya River peak season flow (C. Xu et al., 2019). These uses of $\delta^{18}O$ to develop reconstructions, along with other studies (e.g., Treydte et al., 2006; G. Xu et al., 2019), serve to illustrate that tree ring $\delta^{18}O$ has indeed moved beyond "potential" (Gagen et al., 2011), and earned its place as a practical, valuable paleoclimate proxy.

2.4. Proxy–Streamflow Correlations

As a preliminary investigation, we performed correlation analyses between streamflow and proxy data. Correlations are calculated at different lags: $l = -2$ to +2 years for ring width (Meko et al., 2007), and $l = 0$ to

Table 2
Metadata of δ^{18} O Chronologies

Site	Longitude	Latitude	Species	References
Mae Hong Son	98.93	19.28	*Pinus merkusii*	C. Xu et al. (2015)
Mu Cang Chai	104.06	21.40	*Fokienia hodginsii*	Sano et al. (2012)
Phou Khao Khouay	102.79	18.32	*Pinus merkusii*	C. Xu et al. (2019)
Umpang	98.87	16.09	*Pinus merkusii*	C. Xu et al. (2018)

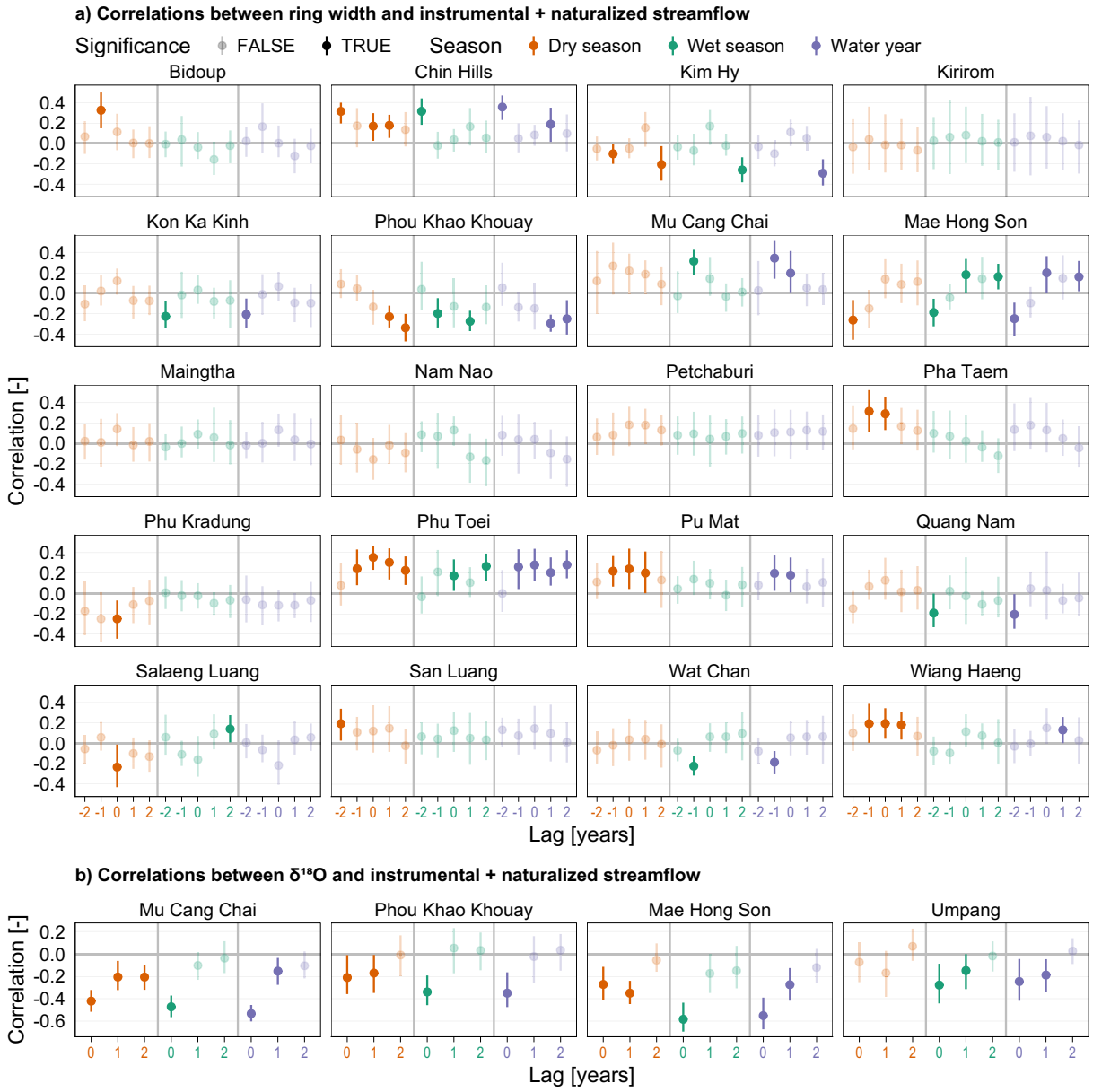

Figure 2. Streamflow–proxy correlations. The error bars show the 5th–95th bootstrapped empirical quantiles obtained from 1,000 replicates, using the stationary bootstrap (Politis & Romano, 1994). The dots indicate the medians. Lag l ldenotes correlations between proxy at year t and streamflow at year $t + l$. Only positive lags are calculated for δ^{18}O.

2 years for δ^{18}O. Negative lags are calculated only for ring with to account for the case when trees use stored carbon from previous years (Fritts, 1976; Stockton & Jacoby, 1976); these processes are not relevant for δ^{18}O, and first-order autocorrelation for δ^{18}O is low (C. Xu et al., 2015). Positive lags account for the case when the catchment's runoff processes are slower than precipitation inputs (Hawley, 1937; Stockton & Jacoby, 1976). For robustness, we repeated the correlation analysis 1,000 times using the stationary bootstrap (Politis & Romano, 1994). The correlation patterns (how the median and confidence intervals of the correlation coefficient change over different seasons at different lags) are shown in Figure 2.

Among the ring width sites, there are multiple correlation patterns (Figure 2a): some sites such as Chin Hills and Phu Toei correlate positively, while others, for example, Phou Khao Khouay (PKK), correlate negatively. The strong negative correlations at PKK is consistent with Buckley, Duangsathaporn, et al. (2007),

who reported significant negative correlations between PKK tree rings and April–June rainfall of the prior and current year. They interpreted this as the trees' responses to light availability. *Pinus merkusii*, like most Pinus species, requires abundant light at the start of the growth season. Heavy rain in this period would increase cloud cover and reduce growth. April–June is also the "shoulder" period between wet and dry seasons, leading to significant negative correlations with streamflow for both seasons, and hence for the water year. The Mae Hong Son site displays a more peculiar correlation pattern: significant negative correlation at $l = -2$ but significant positive correlations at $l = 0$ and $l = 2$. Four sites do not correlate with streamflow at all. These various patterns suggest that the ring width–streamflow relationship is complex and "noisy," Our tree ring sites lie outside the drainage basin of P.1, that is, we rely on teleconnections, on the basis that both streamflow and tree rings are influenced by large scale climate drivers such as the monsoon and ENSO. However, each chronology is also influenced by local site conditions that are not relevant to P.1. These signals are not noise *per se* but they are noise for the task of streamflow reconstructions at P.1. A large number of sites are thus required to extract the strongest common climate signals.

Unlike ring width, δ^{18}O displays more consistent correlation patterns (Figure 2b): all significant correlations are negative, and the strongest correlations are often observed at $l = 0$. Some correlations have magnitudes larger than 0.5, while the largest correlation magnitude in ring width is only 0.36. These observations corroborate that δ^{18}O chronologies may contain stronger climate signals than do ring width chronologies (Gagen et al., 2011; C. Xu et al., 2019).

In general, we observe that ring width tends to correlate more strongly with dry season flow than with wet season flow. Conversely, δ^{18}O tends to correlate more strongly with wet season flow than with dry season flow. Both proxies correlate well with annual flow. The proxy-streamflow correlations observed here are also in agreement with the proxy-precipitation correlation analysis (Text S5). Both analyses show that tree ring proxies have different strength and sensitivity to different parts of the hydrograph, and have the potential to be combined for better seasonal reconstructions.

3. Reconstruction Framework

To build reconstruction models, we define an *input* as a chronology–lag combination that significantly correlates with streamflow. For instance, some inputs for the annual reconstruction are Chin Hills ring width at lag −2, and Umpang δ^{18}O at lag 0 (Figure 2). At P.1, there are 19 inputs for the dry season, 28 for the wet season, and 30 for the water year.

The reconstruction framework consists of two main modules: Regression and Input Selection. In Regression (Section 3.1), the selected inputs for each target are given, and we find the regression coefficients that best match the targets while accounting for mass balance simultaneously, using a *penalized least squares* formulation. In Input Selection (Section 3.2), we find the best subset of inputs that minimizes the penalized least squares. The two modules are unified in a nested optimization framework that includes a rigorous cross-validation scheme to assess reconstruction skills (Section 3.3).

3.1. Mass Balance-Adjusted Regression

Assume that we have a matrix $\mathbf{U_d}$ whose columns contain the selected inputs for the dry season. We first need to remove multicollinearity within $\mathbf{U_d}$. A common approach in dendrohydrology is to perform Principal Component Analysis (PCA) on $\mathbf{U_d}$, then reduce the set of principal components (PCs) to a parsimonious subset that is most relevant to the streamflow target (Coulthard et al., 2016; Hidalgo et al., 2000). Here, we use a backward stepwise PC selection routine (Woodhouse et al., 2006). This transformation from the selected inputs to the selected PCs X_d is denoted as the function $g(.)$:

$$\mathbf{X_d} = g(\mathbf{U_d}) \tag{1}$$

Similarly, given the selected inputs $\mathbf{U_w}$ for the wet season and $\mathbf{U_q}$ for the water year, we apply $g(.)$ to get

$$\mathbf{X_w} = g(\mathbf{U_w}) \tag{2}$$

$$\mathbf{X_q} = g(\mathbf{U_q}) \tag{3}$$

Now, let $\mathbf{y_d}, \mathbf{y_w}$, and $\mathbf{y_q}$ be the target time series of dry season, wet season, and annual streamflow, respectively (these targets can be log-transformed when necessary). Reconstructing streamflow for the three targets means solving the following regression equations:

$$\mathbf{y_d} = \mathbf{X_d}\boldsymbol{\beta}_d + \boldsymbol{\varepsilon}_d \tag{4}$$

$$\mathbf{y_w} = \mathbf{X_w}\boldsymbol{\beta}_w + \boldsymbol{\varepsilon}_w \tag{5}$$

$$\mathbf{y_q} = \mathbf{X_q}\boldsymbol{\beta}_q + \boldsymbol{\varepsilon}_q \tag{6}$$

where $\boldsymbol{\beta}_d, \boldsymbol{\beta}_w$, and $\boldsymbol{\beta}_q$ are the corresponding regression coefficients; and $\boldsymbol{\varepsilon}_d, \boldsymbol{\varepsilon}_w$, and $\boldsymbol{\varepsilon}_q$ are white noise.

Next, let

$$\mathbf{y} = \begin{bmatrix} \mathbf{y_d} \\ \mathbf{y_w} \\ \mathbf{y_q} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{X_d} & & \\ & \mathbf{X_w} & \\ & & \mathbf{X_q} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_d \\ \boldsymbol{\beta}_w \\ \boldsymbol{\beta}_q \end{bmatrix}, \quad \text{and} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_d \\ \boldsymbol{\varepsilon}_w \\ \boldsymbol{\varepsilon}_q \end{bmatrix}. \tag{7}$$

Equations 4–6 can then be converted to a more compact form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{8}$$

Equation 8 has the canonical form of linear regression. It can be solved as a least squares problem:

$$\min_{\boldsymbol{\beta}} \quad J_1 = \left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)' (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \tag{9}$$

yielding the solution

$$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{10}$$

Solving Equation 8 is equivalent to solving Equations 4–6 simultaneously. The three regression problems in Equations 4–6 are independent of one another, and the above formulation places no constraints to match the sum of the seasonal flows to the annual flow. Therefore, such formulation can yield large differences in the annual mass balance. As we shall see later, this happens at station P.1.

To account for mass balance, it is tempting to impose a constraint,

$$\mathbf{X_d}\boldsymbol{\beta}_d + \mathbf{X_w}\boldsymbol{\beta}_w = \mathbf{X_q}\boldsymbol{\beta}_q. \tag{11}$$

But, Equation 11 is often overdetermined: it is a system of T equations, one for each year, and we almost always have more equations than unknowns in a regression problem. Instead, we can add to the objective function in Equation 9 a penalty term that is based on the differences ($\boldsymbol{\delta}$) between the LHS and the RHS of Equation 11.

$$\boldsymbol{\delta} = \mathbf{X_d}\boldsymbol{\beta}_d + \mathbf{X_w}\boldsymbol{\beta}_w - \mathbf{X_q}\boldsymbol{\beta}_q. \tag{12}$$

If the reconstructions involve log-transformed flows, the mass difference is

$$\delta_t = \log\left(\exp\left(\mathbf{x}_{\mathbf{d},t}\boldsymbol{\beta}_d\right) + \exp\left(\mathbf{x}_{\mathbf{w},t}\boldsymbol{\beta}_w\right)\right) - \mathbf{x}_{\mathbf{q},t}\boldsymbol{\beta}_q \qquad \forall t = 1, \ldots, T. \tag{13}$$

Just as we minimize the squared differences between prediction and observation, we also minimize the squared mass differences. Finally, we add a weight λ to represent the importance of the penalty term, and obtain a new objective function

$$\min_{\boldsymbol{\beta}} \quad J_2 = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \lambda\boldsymbol{\delta}'\boldsymbol{\delta} \tag{14}$$

We call this the *penalized least squares* problem. Observe that when $\lambda = 0$, the penalty term disappears, and the penalized least squares problem becomes the canonical least squares problem. The higher λ is, the more important the penalty becomes. How to choose an appropriate λ? Theoretically, an analyst can test different values of λ (a sensitivity analysis) and choose one that produces the best reconstruction skills. Practically,

however, such a task may not be straightforward (e.g., which skill metric to use?), and the analyst may also have her subjective priority on the mass balance criterion. A reasonable approach, then, is to use the sensitivity analysis as a guide to make a more informed (maybe still subjective) decision. We will demonstrate how reconstruction skills vary with λ in Section 4.2, and provide some further guidelines on how to choose λ in Section 4.4. For now, let us assume that an appropriate λ is chosen.

Without flow transformation, $\boldsymbol{\delta}$ is linear (Equation 12), so J_2 is quadratic. We can solve Equation 14 analytically to get

$$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X} + \lambda\mathbf{A}'\mathbf{A})^{-1}\mathbf{X}'\mathbf{y} \tag{15}$$

where $\mathbf{A} = [\mathbf{X_d} \quad \mathbf{X_w} \quad -\mathbf{X_q}]$. The proof is provided in Text S7.

When log-transformations are involved, $\boldsymbol{\delta}$ is not linear, and Equation 14 cannot be solved analytically. But it can be solved numerically using any nonlinear solver. Here, we use an efficient quasi-Newton method called L-BFGS-B (Byrd et al., 1995), available in the R function optim(). We have implemented the mass balance-adjusted regression procedure in the R package mbr (Nguyen, 2021).

3.2. Optimal Input Selection

A consolidated approach to input selection in linear regression problems is to use Branch and Bound algorithms, such as Leaps and Bounds (Furnival & Wilson, 1974) or its more recent variants (Duarte Silva, 2001, 2002). These algorithms are conceived to balance goodness-of-fit with model simplicity. In this work however, we also need to account for mass balance besides goodness-of-fit. Therefore, the input selection routine must explicitly account for the penalized least squares objective (Equation 14). If the number of inputs is small, we can exhaustively search all possible subsets and choose the one that yields the minimum penalized least squares value (PLSV). However, this method quickly becomes infeasible with increasing input size: there are 2^n subsets of n inputs (for station P.1, $n = 19$, 28, and 30). A computationally tractable optimization is necessary (Galelli et al., 2014).

We formulate input selection as a binary optimization problem. Each input has an index, and a binary vector $\mathbf{p}$ encodes input selection: $p_i = 1$ means the ith input is selected. For any given $\mathbf{p}$, that is., for any given input subset, we can solve the mass balance-adjusted regression problem to obtain a PLSV. Our goal then is to find $\mathbf{p}$ that has the best PLSV over all $\mathbf{p}$'s.

Note that $\mathbf{p}$ has three components: $\mathbf{p} = [\mathbf{d} \quad \mathbf{w} \quad \mathbf{q}]'$. Component $\mathbf{d}$ represents the dry season:

$$d_i = \begin{cases} 1 & \text{if proxy } i \text{ is used for the dry season} \\ 0 & \text{otherwise} \end{cases} \qquad i = 1,\ldots,n_d. \tag{16}$$

So, where $d_i = 1$, we take the ith inputs and place into the matrix $\mathbf{U_d}$. Similarly, we create $\mathbf{U_w}$ from $\mathbf{w}$ and $\mathbf{U_q}$ from $\mathbf{q}$. Once we have $\mathbf{U_d}$, $\mathbf{U_w}$, and $\mathbf{U_q}$, the mass balance-adjusted regression procedure can be applied. To improve the robustness of the input selection, the regression is cross-validated 50 times (Section 3.3), each yielding one PLSV estimate. The average of all runs, denoted $f(\mathbf{p})$, is used as the final PLSV for $\mathbf{p}$.

The remaining task is to solve

$$\min_{\mathbf{p}} f(\mathbf{p}). \tag{17}$$

We solve Equation 17 with Genetic Algorithm (Holland, 1975), a metaheuristic optimization technique that allows us to treat the underlying regression as a black-box while searching for the best subset of inputs (Kohavi & John, 1997), and is well suited for binary optimization (Whitley, 1994). We use the R package GA (Scrucca, 2013). Details about the implementations are provided in Text S8.

3.3. Model Assessment

To explore the effect of the mass balance adjustment procedure, we build reconstruction models for $\lambda \in [0, 3]$ with increments of 0.2. Other than the different values for λ, all models are trained exactly the same way, following Sections 3.1 and 3.2.

During optimization, multiple reconstructions are created while the optimal $\mathbf{p}$ is sought for each model. These reconstructions are assessed with the PLSV. The final reconstructions, created with the optimal inputs, are further assessed *post hoc* with the commonly used metrics: coefficient of determination (R^2), reduction of error (RE), and coefficient of efficiency (CE) (Fritts, 1976; Nash & Sutcliffe, 1970). All metrics are calculated over 50 cross-validation runs. In each run, a contiguous chunk of 25% of the data (21 years) is held out, and the model is calibrated on the remaining 75%. R^2 is calculated on the calibration chunks while RE and CE are calculated on the validation chunks. The contiguous chunks aim to test whether the reconstruction can capture regime shifts in the time series, in line with the traditional split-sample scheme; meanwhile, the 50 repetitions provide a distribution for each skill metric, allowing more robust estimation of the mean skill score (Nguyen et al., 2020). More importantly, the distributions enable us to compare skills statistically among models. We use the Wilcoxon rank sum test (Wilcoxon, 1945), also known as the Mann-Whitney test (Mann & Whitney, 1947), to determine whether any skill improvement or degradation from using a positive λ value, as opposed to using $\lambda = 0$, is statistically significant. The Wilcoxon test is chosen because it is nonparametric and does not assume normality.

Finally, we assess the validity of the input selection scheme by examining the weight of each input in the final reconstruction models. The weight w_i of input i is the inner product of its loadings $\mathbf{z}_i$ on the PCs and the PCs' regression parameters $\boldsymbol{\beta}^+$ ($\boldsymbol{\beta}$ without the intercept terms). Mathematically,

$$w_i = \mathbf{z}_i \boldsymbol{\beta}^+ = \sum_{j \in S} z_{ij} \beta_j \tag{18}$$

where S is the index set of selected PCs, z_{ij} is the loading of input i on PC j, and $\boldsymbol{\beta}_j$ is PC j's regression coefficient.

4. Results

4.1. Reconstruction Skills as Functions of λ

Herein, the reconstruction model obtained with $\lambda = 0$ (without mass balance adjustment) is referred to as Model 0. We first compare the skill scores of Model 0 with those of each positive-λ model (i.e., with mass balance adjustment). This comparison is shown in Figure 3. Except for the wet season's R^2, the mean scores of all metrics either are improved or stayed about the same when using positive λ values. Improvements are most consistent with $\lambda \in [0.6, 2]$. However, the improved scores are often not statistically significant under the Wilcoxon test ($\alpha = 0.1$), so we do not claim that the mass balance-adjusted regression improve skills necessarily. But it is reassuring that the new method does not worsen skills while attempting to account for the annual mass balance.

To understand how the mass balance adjustment can improve skills, let us recall the models' formulation. Model 0 reconstructs the dry season, wet season, and annual flows independently. Contrarily, the mass balance adjustment procedure links all three reconstructions together via the penalty term (Equations 12–14). This link provides each reconstruction with additional information from the other two, thereby improving skills.

Wet season's R^2 is a curious exception. Recall that R^2 measures goodness-of-fit while RE and CE measures generalizability. The mass balance regression traded off goodness-of-fit for better out-of-sample prediction. This trade-off should reduce overfitting and result in more robust reconstructions. However, the trade-off is not entirely successful: the decreases in R^2 are statistically significant while the increases in RE and CE are not. This trade-off shows that the wet season is the most difficult target to reconstruct; such a trade-off did not occur for the dry season and the water year. We shall revisit this issue in Section 4.2.

Figure 3 suggests that $\lambda \in [0.6, 2]$ is a reasonable range of values to choose from, based on skills alone. We choose $\lambda = 1.2$ for the subsequent discussion—this is the only value that resulted in significantly higher RE and CE for the water year. Certainly this choice is subjective and somewhat arbitrary; any $\lambda \in [0.6, 2]$ would be a reasonable choice. Henceforth, the model based on $\lambda = 1.2$ is referred to as Model 1. We now compare Model 0's and Model 1's reconstructions in more detail.

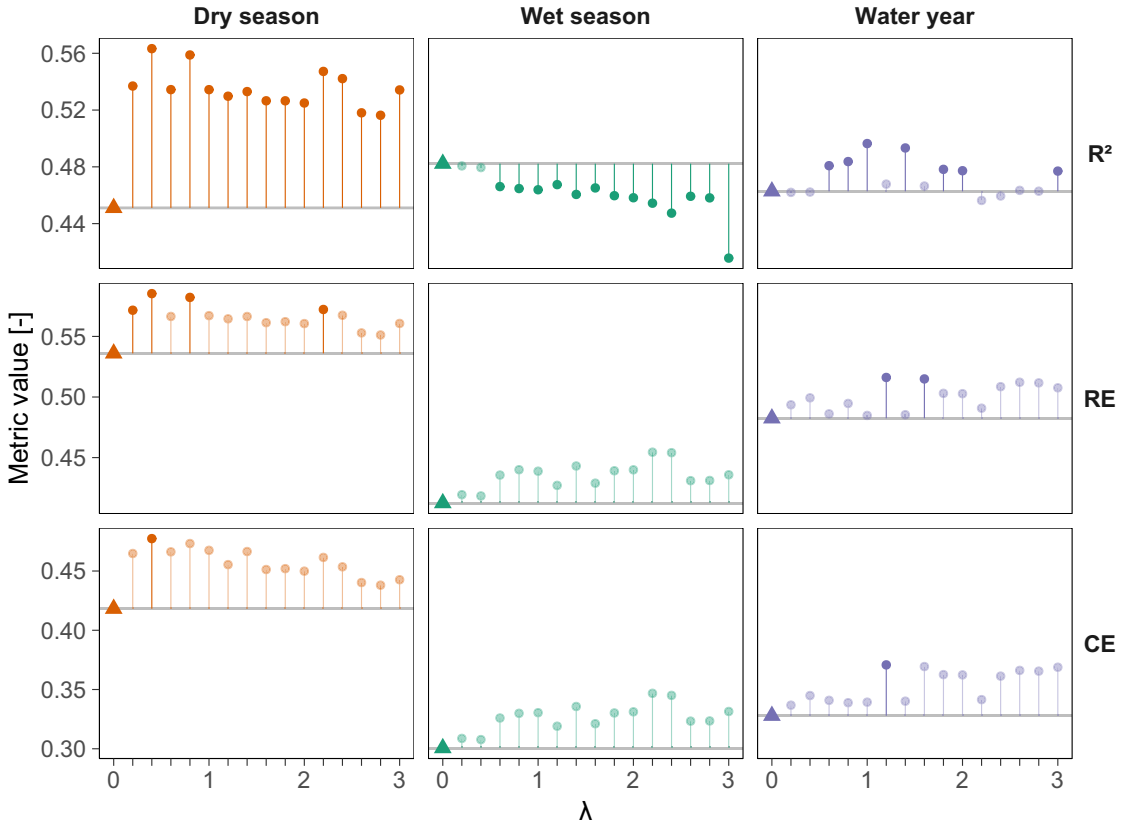

Figure 3. Mean reconstruction skill scores resulted from changing the mass balance penalty weight λ in Equation 14. The scores from $\lambda = 0$ are marked with triangles and gray horizontal lines for easy comparison. Scores from positive λ values that are not significantly different from $\lambda = 0$ under a Wilcoxon rank sum test ($\alpha = 0.1$) are shown with faint colors. See also Figure S11 for the distributions of skill scores.

4.2. Reconstructions

For both Model 0 and Model 1, the reconstructions match their target closely (Figure 4a). Exceptions are the extreme wet years of 1971, 1973, 1975, and 1978, when consecutive La Niña events intensified monsoon rain, and large tropical storms carried abundant moisture to the basin in the wet season, causing severe floods (Lim & Boochabun, 2012). These extreme events proved difficult for the wet season and annual reconstructions. In these years, the dry season reconstructions are good, implying that the mismatches in the annual reconstructions are due purely to the wet season's flow. During strong monsoon seasons, the soil is fully saturated. The ring width chronologies cannot capture saturation excess flow. While this part of the hydrograph can be recorded by $\delta^{18}O$, there are only four $\delta^{18}O$ chronologies in our network, limiting the models' ability to capture extremely high flows. Seeing this helps us understand better our observation earlier that the wet season is the most difficult target. Perhaps we have exhausted the information contained in the proxies for the wet season. Hence, Model 1, although making used of all three seasons, had to trade-off goodness-of-fit for generalizability for this target. Also, there may be nonlinearities in the streamflow–proxy relationships at the extremes. In future studies, nonlinear reconstruction models (e.g., Nguyen & Galelli, 2018; Torbenson & Stagge, 2021) could be incorporated to address this problem.

The reconstructions provide some interesting insights into the inter- and intra-annual variability of the river (Figure 4b). Between 1825 and 1855, sustained low flow was observed in the wet season and water year reconstructions. However, in the dry season, the low flow period ended 15 years earlier, around 1840. Conversely, a period of sustained high flow was observed in all three reconstructions between 1790 and 1820, especially for the dry season. Most notably, dry season flow in 1815 was so high that it accounted for more than 50% of the annual flow—a rare event that occurred in only nine of 254 years (Figure 5). Interestingly, these nine years span the whole range of annual flow in the reconstructions, from 1,000 to 3,000 million m^3.

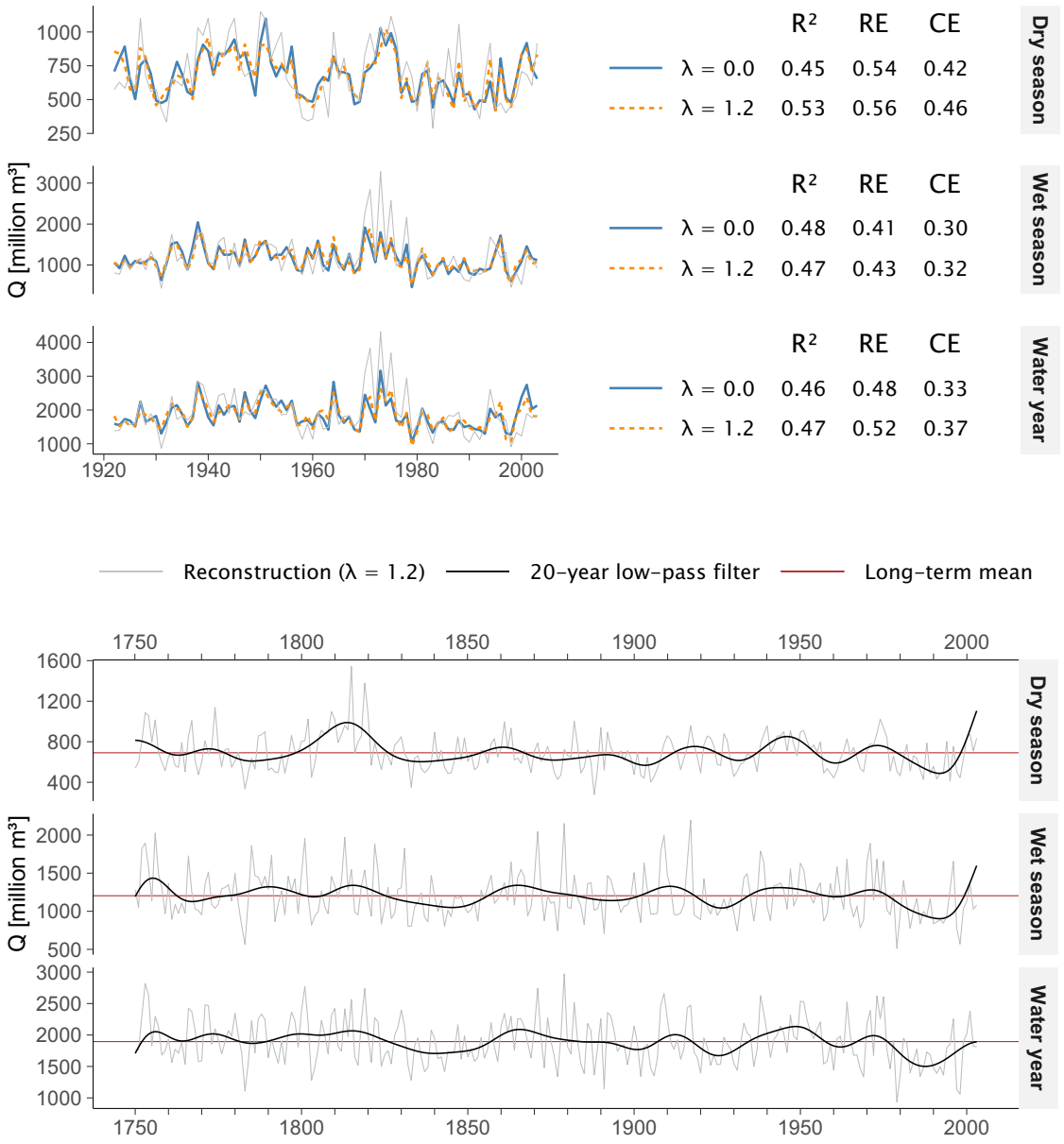

Figure 4. (a) Reconstruction skills and time series for the instrumental period, produced using two models: without mass balance adjustment ($\lambda = 0$) and with adjustment ($\lambda = 1.2$) in the regression problem (Equation 14). Gray lines show naturalized observations. A comparison for the full horizon is also provided in Figure S4. (b) Full reconstructions with $\lambda = 1.2$.

Statistical desegregation methods, with their fixed assumptions of the sub-annual to annual flow ratios or quantiles, are unlikely to reproduce this wide range of intra-annual variability. Revealing events like these is of fundamental importance for the operations of Bhumibol Dam, which supports large irrigation districts.

The years 1971, 1973, and 1975 once again stand out: extremely high flow occurred in the wet season, leading to small dry season to annual flow (D/Q) ratios (Figure 5). As discussed in Section 4.2, these extreme high flows could not be captured by our proxies. Apart from those three extreme years, the distributions of D/Q are similar in the instrumental record and in the reconstruction.

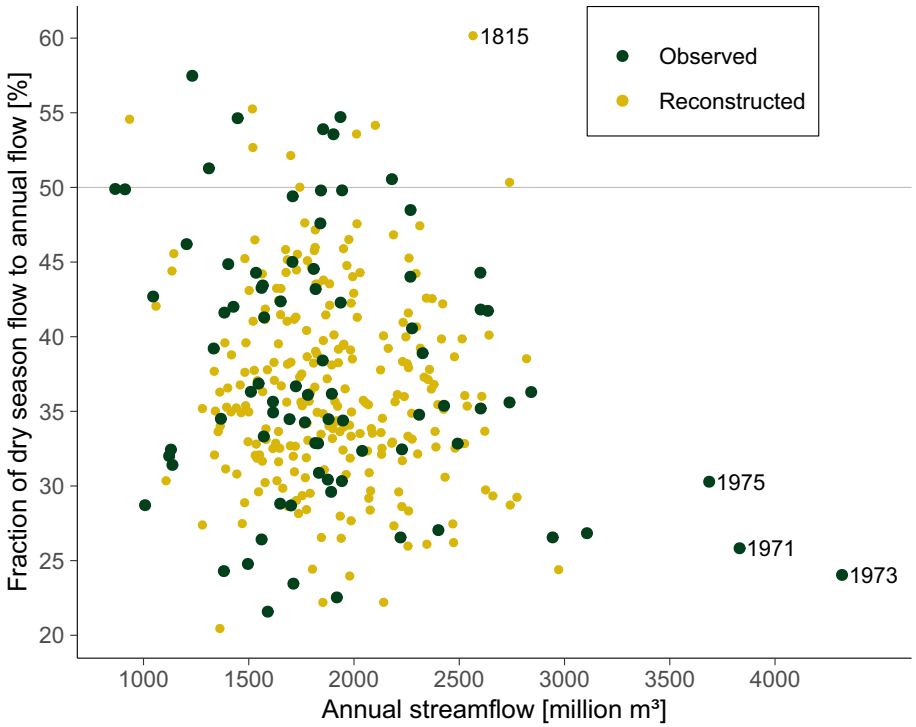

Figure 5. Distribution of the fraction of dry season flow to total annual flow, plot against the annual flow volume. Some extreme years are annotated.

4.3. Selected Inputs

The selected input subsets by both models provide further insights into their similarities and differences. Both models share many common inputs (Figure 6), especially the $\delta^{18}O$ chronologies at Mae Hong Son and Mu Cang Chai, and the ring width chronologies at Phou Khao Khouay. We were surprised that the Mae Hong Son and Mu Cang Chai $\delta^{18}O$ were selected consistently for the dry season, with strong weights placed on the Mu Chang Chai. Upon closer examinations of the correlation analysis (Figure 2), the puzzle is solved. While these $\delta^{18}O$ chronologies correlate less with dry season flow than they do with wet season flow, the correlations are still stronger than those observed at many ring width sites.

For most inputs, the signs of the weights in Figure 6 are the same as the signs of correlation in Figure 2. There are only three exceptions in each model; in these cases the correlations are relatively weak and the weights are near zero. These observations further support that the input selection and the model identification process work generally well and are quite robust.

4.4. Annual Mass Balance

We have now established that the mass balance adjustment produced a sensible model with valid results that are slightly better than, if not comparable to, the conventional model. Finally, we can examine the effect of this new method on the annual mass balance, its ultimate goal.

For each model, we calculate the mass difference, ΔQ, between the total seasonal flow and the annual flow, then examine its trajectory and distribution (Figures 7a and 7b). The mass difference for Model 0 ranges from –896 million m^3 to 932 million m^3, while that range for Model 1 is –432 million m^3 to 748 million m^3; a 45% reduction in range. Moreover, Model 0 yields a mass difference outside the interval ±190 million m^3 (±10% of the mean annual flow (MAF); shaded region in Figure 7b) in 48% of the years. That figure for Model 1 is only 21%. By these metrics, Model 1 is much better than Model 0 in preserving mass balance.

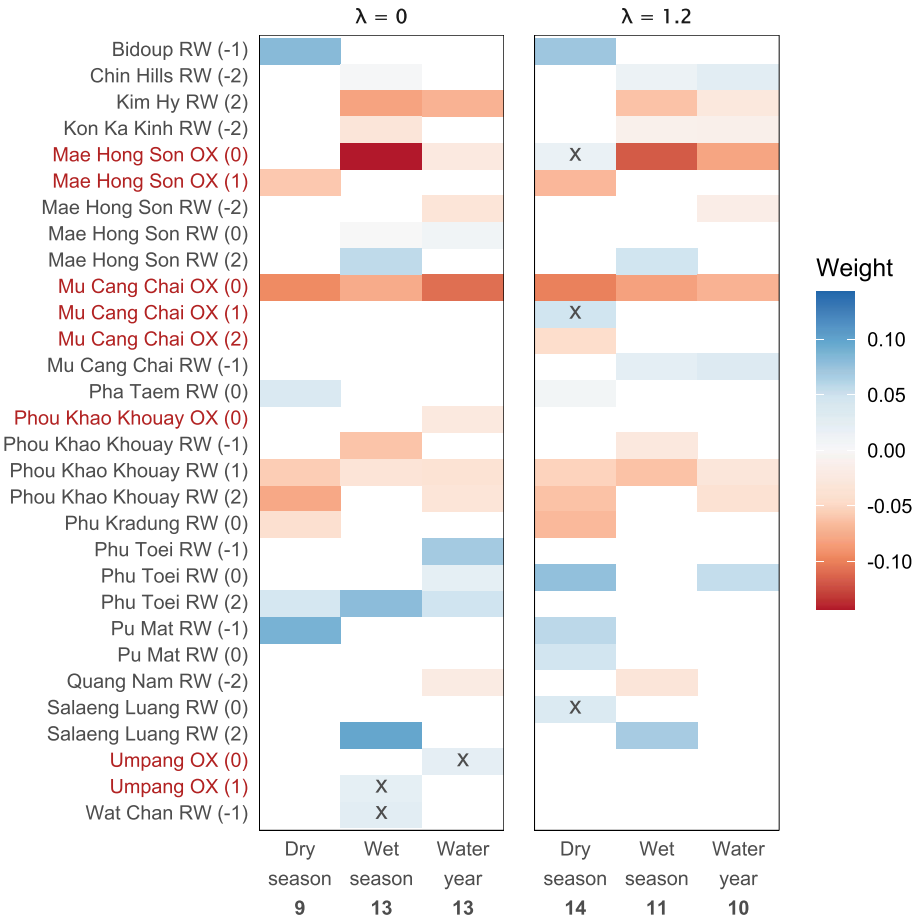

Figure 6. The weights of selected inputs for each streamflow target (columns) in each model (panels). Only inputs that were selected at least once are shown. The stable oxygen isotope (OX) inputs are marked in red. Inputs whose weights have different signs from their correlation coefficients are marked with "X." Bold numbers at the bottom row of the *x*-axis are the number of selected inputs in each column.

As λ increases, the distribution of ΔQ tends to narrow (Figure 7c) and the percentage of years with ΔQ outside 10% MAF generally decreases. The fluctuations in the decreasing trend are likely due to GA being stuck in local optima in some cases—a common challenge with metaheuristic search techniques. Despite this, the decreasing trend is clear. With $\lambda = 3$, the range of ΔQ is narrowed down to [-403, 511] million m^3, and the percentage of ΔQ outside 10% MAF is 12%. The choice of $\lambda = 3$ addresses the mass balance objective the best; for an analyst who is most concerned about mass balance this value may be a better choice than $\lambda = 1.2$. However, $\lambda = 3$ resulted in the largest decrease in wet season's R^2. Ultimately, the choice of λ depends on the analyst's priorities (skills vs. mass balance).

5. Discussion

The "typical" format of streamflow reconstructions hinders their operational usability: reconstructions often target a specific component of the hydrograph and offer a temporal resolution that is too coarse for water management decisions. Our results show that both challenges could be tackled by leveraging the seasonal sensitivities of different proxies, a modeling approach that enables skillful seasonal (and annual) reconstructions for the inflow to Bhumibol Dam. These reconstructions may be of great value to water managers when seen in light of the current decision-making framework. There are indeed two pivotal moments in the annual management of the Chao Phraya River basin: the dry and wet season water allocations, through which water is allocated to the basin's sectors and regions (Divakar et al., 2011; Takeda et al., 2016). The allocation builds on a few different criteria, including the amount of water stored in Bhumibol and

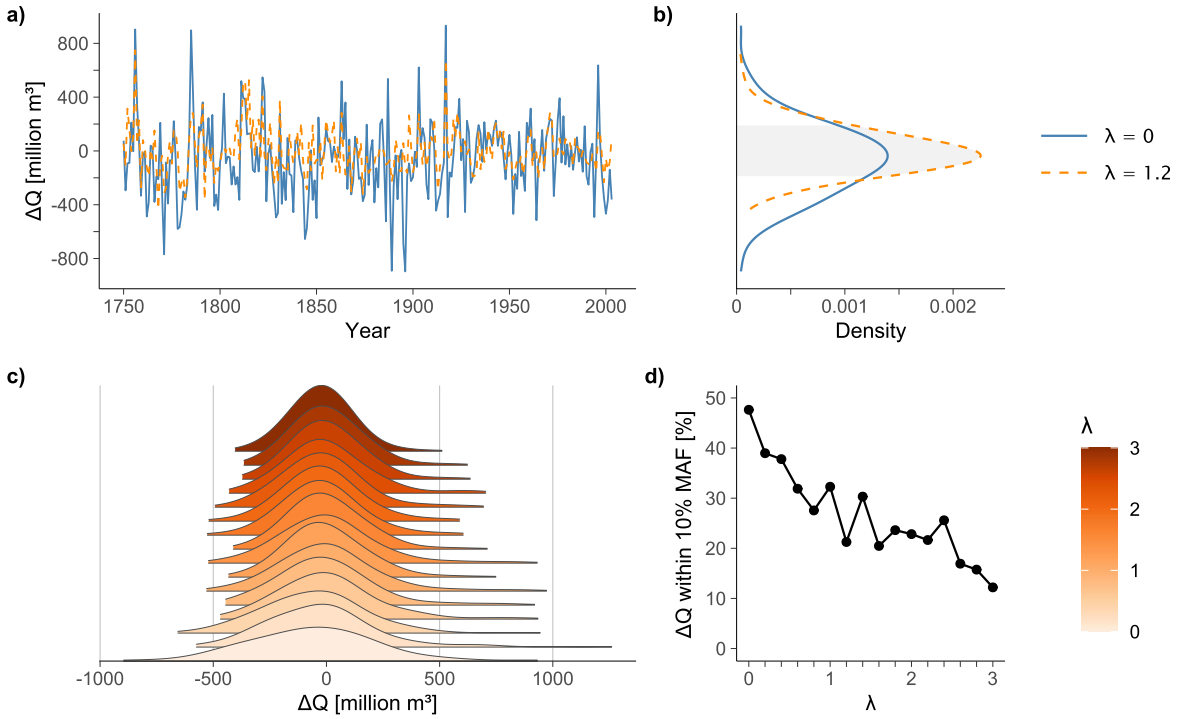

Figure 7. (a) The differences, ΔQ, between the total seasonal flow and the annua flow. (b) Distributions of ΔQ; the shaded region denotes the ±190 million m^3 range, equivalent to ±10% of the mean annual flow (MAF). (c) Distribution of ΔQ as a function of λ. (d) Percentage of ΔQ within 10% MAF as a function of λ.

Sirikit dams as well as a projection of seasonal water demand, which in turn depends on the expected water availability. Our reconstructions show that the dry and wet season flows can deviate from the observed envelope of variability—for example, there are instances in which the dry season flow accounts for more than 50% of the annual flow—meaning that the hydro-climatological risk to which water users are exposed is not properly characterized. Such information could be used to evaluate the robustness of the current allocation process or to stress-test alternative allocation schemes (Divakar et al., 2011).

A second fundamental aspect of streamflow reconstructions is accuracy. In our analysis, we showed that the optimal input selection procedure yields good reconstruction skills for both model setups, and that Model 1, by imposing a mass balance adjustment, produces more robust reconstructions than does Model 0 (Section 4.2). More importantly, the adjustment significantly reduces the differences between the total seasonal flow and the annual flow (Section 4.4). Without the adjustment, the mass difference can be as large as 932 million m^3, or about 49% of the mean annual flow. It amounts to 130% of the irrigation demand from the Ping River downstream of Bhumibol Reservoir (Divakar et al., 2011). Such a discrepancy would arguably restrict the use of reconstructions in the aforementioned decision-making context. But with the adjustment, both the frequency and magnitude of discrepancies are reduced: this is crucial for making reconstructions appealing to water managers. Therefore, the mass balanced-adjusted regression produces more reliable streamflow distributions, which then can be used in bottom-up vulnerability assessments, a popular form of analysis in the water resources domain (Borgomeo, Farmer, & Hall, 2015; Borgomeo, Pflug, et al., 2015; Herman et al., 2016; Pielke et al., 2012).

The modeling framework that we proposed here can be reapplied and expanded in several ways. First, the mass balance formulation is applicable to other climate variables such as precipitation, and to other contexts where a penalty term in the regression equation is desirable. For example, if one wishes to reconstruct streamflow at two tributaries as well as the main stream of a river, the mass balance adjustment should be used to minimize the differences between the total flow of the tributaries and the flow on the main stream. Second, if more proxies—such as wood density (Schweingruber & Briffa, 1996) or blue intensity (Buckley et al., 2018)—are available, they can also be incorporated into the reconstruction framework. Finally, the mass balance formulation is extendable to higher resolutions, for example, quarterly or monthly, as long as

the proxy network is sensitive enough to the higher resolution targets. These directions can help dendrohydrology realize its value in operational water management, an area where annual, unconstrained streamflow reconstructions have had limited success.

6. Conclusions

In almost a century-long history, dendrohydrology has largely contributed to our understanding of global hydrological variability, as well as its linkages to large scale climate features. But to really fulfill their potential, streamflow reconstructions should become a "tool" regularly available to water managers. That means improving their format and accuracy. Our work shows that advanced data analytics, informed by multi-proxy networks, represent a technical pathway toward meeting these grand challenges.

Data Availability Statement

Streamflow data were obtained from the Thai Royal Irrigation Department at http://hydro-1.rid.go.th/index.html. The authors provide all data, documented code, and results in an online repository at http://github.com/ntthung/multiproxy?mbr (https://doi.org/10.5281/zenodo.4306343).

Acknowledgments
The authors are indebted to Edward Cook for the updated chronologies (Salaeng Luang, Mae Hong Son, and Mu Cang Chai), and for his valuable comments. The authors are also grateful to Thanh Dang for the VIC-Res model output. The diligent work of Le Canh Nam and many other local colleagues in collecting tree core samples over many years is crucial in building the tree ring network, and is sincerely appreciated. The authors thank Adam Csank for the historical background of dendrohydrology through his AAG 2019 talk and subsequent conversation. Hung Nguyen is supported by the President's Graduate Fellowship from the Singapore University of Technology and Design. Chenxi Xu is supported by the Chinese Academy of Sciences Pioneer Hundred Talents Program, the National Natural Science Foundation of China (No. 42022059 and 41888101), and the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDB26020000). Brendan Buckley is supported by the US National Science Foundation grants No. AGS-1602629 and AGS-2001949.

References

Borgomeo, E., Farmer, C. L., & Hall, J. W. (2015). Numerical rivers: A synthetic streamflow generator for water resources vulnerability assessments. *Water Resources Research*, *51*(7), 5382–5405. https://doi.org/10.1002/2014WR016827

Borgomeo, E., Pflug, G., Hall, J. W., & Hochrainer-Stigler, S. (2015). Assessing water resource system vulnerability to unprecedented hydrological drought using copulas to characterize drought duration and deficit. *Water Resources Research*, *51*(11), 8927–8948. https://doi.org/10.1002/2015WR017324

Buckley, B. M., Anchukaitis, K. J., Penny, D., Fletcher, R., Cook, E. R., Sano, M., et al. (2010). Climate as a contributing factor in the demise of Angkor, Cambodia. *Proceedings of the National Academy of Sciences*, *107*(15), 6748–6752. https://doi.org/10.1073/pnas.0910827107

Buckley, B. M., Barbetti, M., Watanasak, M., D'Arrigo, R., Boonchirdchoo, S., & Sarutanon, S. (1995). Dendrochronological investigations in Thailand. *IAWA Journal*, *16*(4), 393–409. https://doi.org/10.1163/22941932-90001429

Buckley, B. M., Duangsathaporn, K., Palakit, K., Butler, S., Syhapanya, V., & Xaybouangeun, N. (2007). Analyses of growth rings of *Pinus merkusii* from Lao P.D.R. *Forest Ecology and Management*, *253*(1), 120–127. https://doi.org/10.1016/j.foreco.2007.07.018

Buckley, B. M., Fletcher, R., Wang, S.-Y. S., Zottoli, B., & Pottier, C. (2014). Monsoon extremes and society over the past millennium on mainland Southeast Asia. *Quaternary Science Reviews*, *95*, 1–19. https://doi.org/10.1016/j.quascirev.2014.04.022

Buckley, B. M., Hansen, K. G., Griffin, K. L., Schmiege, S., Oelkers, R., D'Arrigo, R. D., et al. (2018). Blue intensity from a tropical conifer's annual rings for climate reconstruction: An ecophysiological perspective. *Dendrochronologia*, *50*, 10–22. https://doi.org/10.1016/j.dendro.2018.04.003

Buckley, B. M., Palakit, K., Duangsathaporn, K., Sanguantham, P., & Prasomsin, P. (2007). Decadal scale droughts over northwestern Thailand over the past 448 years: Links to the tropical Pacific and Indian Ocean sectors. *Climate Dynamics*, *29*(1), 63–71. https://doi.org/10.1007/s00382-007-0225-1

Buckley, B. M., Stahle, D. K., Luu, H. T., Wang, S. Y. S., Nguyen, T. Q. T., Thomas, P., et al. (2017). Central Vietnam climate over the past five centuries from cypress tree rings. *Climate Dynamics*, *48*(11–12), 3707–3723. https://doi.org/10.1007/s00382-016-3297-y

Buckley, B. M., Ummenhofer, C. C., D'Arrigo, R. D., Hansen, K. G., Truong, L. H., Le, C. N., & Stahle, D. K. (2019). Interdecadal Pacific Oscillation reconstructed from trans-Pacific tree rings: 1350–2004 CE. *Climate Dynamics*, *53*(5–6), 3181–3196. https://doi.org/10.1007/s00382-019-04694-4

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, *16*(5), 1190–1208. https://doi.org/10.1137/0916069

Chowdhury, A. F. M. K., Dang, T. D., Nguyen, H. T. T., Koh, R., & Galelli, S. (2021). The Greater Mekong's climate-water-energy nexus: How ENSO-triggered regional droughts affect power supply and CO_2 emissions. *Earth's Future*, *9*(3), e2020EF001814. https://doi.org/10.1029/2020EF001814

Cook, B. I., & Buckley, B. M. (2009). Objective determination of monsoon season onset, withdrawal, and length. *Journal of Geophysical Research*, *114*(D23), D23109. https://doi.org/10.1029/2009JD012795

Cook, E. R., Anchukaitis, K. J., Buckley, B. M., D'Arrigo, R. D., Jacoby, G. C., & Wright, W. E. (2010). Asian monsoon failure and megadrought during the last millennium. *Science*, *328*(5977), 486–489. https://doi.org/10.1126/science.1185188

Coulthard, B., Smith, D. J., & Meko, D. M. (2016). Is worst-case scenario streamflow drought underestimated in British Columbia? A multi-century perspective for the south coast, derived from tree-rings. *Journal of Hydrology*, *534*, 205–218. https://doi.org/10.1016/j.jhydrol.2015.12.030

D'Arrigo, R., Abram, N. J., Ummenhofer, C., Palmer, J. G., & Mudelsee, M. (2011). Reconstructed streamflow for Citarum River, Java, Indonesia: Linkages to tropical climate dynamics. *Climate Dynamics*, *36*(3–4), 451–462. https://doi.org/10.1007/s00382-009-0717-2

D'Arrigo, R., Barbett, M., Watanasak, M., Buckley, B. M., Krusic, P., Boonchirdchoo, S., & Sarutanon, S. (1997). Progress in dendroclimatic studies of mountain pine in northern Thailand. *IAWA Journal*, *18*(4), 433–444. https://doi.org/10.1163/22941932-90001508

D'Arrigo, R., Palmer, J. G., Ummenhofer, C. C., Kyaw, N. N., & Krusic, P. (2011). Three centuries of Myanmar monsoon climate variability inferred from teak tree rings. *Geophysical Research Letters*, *38*(24). https://doi.org/10.1029/2011GL049927

Divakar, L., Babel, M., Perret, S., & Gupta, A. D. (2011). Optimal allocation of bulk water supplies to competing use sectors based on economic criterion – An application to the Chao Phraya River Basin, Thailand. *Journal of Hydrology*, *401*(1–2), 22–35. https://doi.org/10.1016/j.jhydrol.2011.02.003

Duarte Silva, A. P. (2001). Efficient variable screening for multivariate analysis. *Journal of Multivariate Analysis*, *76*(1), 35–62. https://doi.org/10.1006/jmva.2000.1920

Duarte Silva, A. P. (2002). Discarding variables in a principal component analysis: Algorithms for all-subsets comparisons. *Computational Statistics*, *17*(2), 251–271. https://doi.org/10.1007/s001800200105

Electricity Generating Authority of Thailand. (2013). *Exploring EGAT power plants and dams*. Retrieved from https://www.egat.co.th/en/information/power-plants-and-dams

FAO. (2017). *AQUASTAT database*. Retrieved from http://www.fao.org/aquastat/statistics/query/index.html

Fritts, H. C. (1976). *Tree rings and climate*. Elsevier. https://doi.org/10.1016/B978-0-12-268450-0.X5001-0

Fritts, H. C., Blasing, T. J., Hayden, B. P., & Kutzbach, J. E. (1971). Multivariate techniques for specifying tree-growth and climate relationships and for reconstructing anomalies in paleoclimate. *Journal of Applied Meteorology*, *10*(5), 845–864. https://doi.org/10.1175/1520-0450(1971)010⟨0845:MTFSTG⟩2.0.CO;2

Furnival, G. M., & Wilson, R. W. (1974). Regressions by leaps and bounds. *Technometrics*, *16*(4), 499–511. https://doi.org/10.1080/00401706.1974.10489231

Gagen, M., McCarroll, D., Loader, N. J., & Robertson, I. (2011). Stable isotopes in dendroclimatology: Moving beyond 'potential'. In M. K. Hughes, T. W. Swetnam, & H. F. Diaz (Eds.), *Dendroclimatology: Progress and prospects* (pp. 147–172). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5725-0_6

Galelli, S., Humphrey, G. B., Maier, H. R., Castelletti, A., Dandy, G. C., & Gibbs, M. S. (2014). An evaluation framework for input variable selection algorithms for environmental data-driven models. *Environmental Modelling & Software*, *62*, 33–51. https://doi.org/10.1016/j.envsoft.2014.08.015

Galelli, S., Nguyen, H. T. T., Turner, S. W. D., & Buckley, B. M. (2021). Time to use dendrohydrological data in water resources management? *Journal of Water Resources Planning and Management*, *147*(8), 01821001. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001422

Güner, H. T., Köse, N., & Harley, G. L. (2017). A 200-year reconstruction of Kocasu River (Sakarya River Basin, Turkey) streamflow derived from a tree-ring network. *International Journal of Biometeorology*, *61*(3), 427–437. https://doi.org/10.1007/s00484-016-1223-y

Hansen, K. G., Buckley, B. M., Zottoli, B., D'Arrigo, R. D., Nam, L. C., Van Truong, V., et al. (2017). Discrete seasonal hydroclimate reconstructions over northern Vietnam for the past three and a half centuries. *Climatic Change*, *145*(1–2), 177–188. https://doi.org/10.1007/s10584-017-2084-z

Hardman, G., & Reil, O. E. (1936). The relationship between tree-growth and stream runoff in the Truckee River Basin, California-Nevada (Vol. 141). Nevada Agricultural Experiment Station Bulletin.

Hawley, F. M. (1937). Relationship of southern cedar growth to precipitation and run off. *Ecology*, *18*(3), 398–405. https://doi.org/10.2307/1931209

Herman, J. D., Zeff, H. B., Lamontagne, J. R., Reed, P. M., & Characklis, G. W. (2016). Synthetic drought scenario generation to support bottom-up water supply vulnerability assessments. *Journal of Water Resources Planning and Management*, *142*(11), 04016050. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000701

Hidalgo, H. G., Piechota, T. C., & Dracup, J. A. (2000). Alternative principal components regression procedures for dendrohydrologic reconstructions. *Water Resources Research*, *36*(11), 3241–3249. https://doi.org/10.1029/2000WR900097

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. The University of Michigan Press.

Josse, J., & Husson, F. (2016). missMDA: A package for handling missing values in multivariate data analysis. *Journal of Statistical Software*, *70*(1). https://doi.org/10.18637/jss.v070.i01

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, *97*(1–2), 273–324. https://doi.org/10.1016/s0004-3702(97)00043-x

Lieberman, V. (2003). *Strange parallels: Volume 1, Integration on the Mainland: Southeast Asia in Global Context, c.800–1830*. Cambridge University Press.

Lieberman, V., & Buckley, B. M. (2012). The impact of climate on Southeast Asia, circa 950–1820: New Findings. *Modern Asian Studies*. *46*, 1049–1096. https://doi.org/10.1017/S0026749X12000091

Lim, H. S., & Boochabun, K. (2012). Flood generation during the SW monsoon season in northern Thailand. *Geological Society, London, Special Publications*, *361*(1), 7–20. https://doi.org/10.1144/SP361.3

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, *18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Meko, D. M., & Woodhouse, C. A. (2011). Application of streamflow reconstruction to water resources management. In M. K. Hughes, T. W. Swetnam, & H. F. Diaz (Eds.), *Dendroclimatology: Progress and prospects* (pp. 231–261). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5725-0_8

Meko, D. M., Woodhouse, C. A., Baisan, C. A., Knight, T., Lukas, J. J., Hughes, M. K., & Salzer, M. W. (2007). Medieval drought in the upper Colorado River Basin. *Geophysical Research Letters*, *34*(10), L10705. https://doi.org/10.1029/2007GL029988

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Nguyen, H. T. T. (2021). *Mbr: Mass-balance Regression. R package version 0.0.1*.

Nguyen, H. T. T., & Galelli, S. (2018). A linear dynamical systems approach to streamflow reconstruction reveals history of regime shifts in northern Thailand. *Water Resources Research*, *54*(3), 2057–2077. https://doi.org/10.1002/2017WR022114

Nguyen, H. T. T., Turner, S. W. D., Buckley, B. M., & Galelli, S. (2020). Coherent streamflow variability in monsoon Asia over the past eight centuries—Links to oceanic drivers. *Water Resources Research*, *56*(12). https://doi.org/10.1029/2020WR027883

Pielke, R. A., Wilby, R., Niyogi, D., Hossain, F., Dairuku, K., Adegoke, J., et al. (2012). Dealing with complexity and extreme events using a bottom-up, resource-based vulnerability perspective. In *Extreme events and natural hazards: The complexity perspective* (pp. 345–359). American Geophysical Union. https://doi.org/10.1029/2011GM001086

Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, *89*(428), 1303–1313. https://doi.org/10.1080/01621459.1994.10476870

Prairie, J., Nowak, K., Rajagopalan, B., Lall, U., & Fulp, T. (2008). A stochastic nonparametric approach for streamflow generation combining observational and paleoreconstructed data. *Water Resources Research*, *44*(6), 1–11. https://doi.org/10.1029/2007WR006684

Prairie, J., Rajagopalan, B., Lall, U., & Fulp, T. (2007). A stochastic nonparametric technique for space-time disaggregation of streamflows. *Water Resources Research*, *43*(3), 1–10. https://doi.org/10.1029/2005WR004721

Rao, M. P., Cook, E. R., Cook, B. I., D'Arrigo, R., Palmer, J. G., Lall, U., et al. (2020). Seven centuries of reconstructed Brahmaputra River discharge demonstrate underestimated high discharge and flood hazard frequency. *Nature Communications*, *11*(1), 6017. https://doi.org/10.1038/s41467-020-19795-6

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*(8), 5687–5701. https://doi.org/10.1029/2018WR023080

Sano, M., Buckley, B. M., & Sweda, T. (2009). Tree-ring based hydroclimate reconstruction over northern Vietnam from Fokienia hodginsi: Eighteenth century mega-drought and tropical Pacific influence. *Climate Dynamics*, *33*(2–3), 331–340. https://doi.org/10.1007/s00382-008-0454-y

Sano, M., Xu, C., & Nakatsuka, T. (2012). A 300-year Vietnam hydroclimate and ENSO variability record reconstructed from tree ring $\delta^{18}O$. *Journal of Geophysical Research: Atmospheres*, *117*(D12). https://doi.org/10.1029/2012JD017749

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

Schulman, E. (1945). *Tree-ring hydrology of the Colorado River Basin*. University of Arizona Bulletin Series, Laboratory of Tree-Ring Research Bulletin. No. 2, 16(4).

Schweingruber, F. H., & Briffa, K. R. (1996). Tree-ring density networks for climate reconstruction. In P. D. Jones, R. S. Bradley, & J. Jouzel (Eds.), *Climatic variations and forcing mechanisms of the last 2000 years* (pp. 43–66). Springer. https://doi.org/10.1007/978-3-642-61113-1_3

Scrucca, L. (2013). GA: A package for genetic algorithms in R. *Journal of Statistical Software*, *53*(4), 1–37. https://doi.org/10.18637/jss.v053.i04

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

Stockton, C. W. (1971). *The feasibility of augmenting hydrologic records using tree-ring data* (PhD Thesis). University of Arizona.

Stockton, C. W., & Jacoby, G. C. (1976). Long-term surface-water supply and streamflow trends in the Upper Colorado River Basin based on tree-ring analyses. *Lake Powell Research Project Bulletin*, *18*.

Takeda, M., Laphimsing, A., & Putthividhya, A. (2016). Dry season water allocation in the Chao Phraya River basin, Thailand. *International Journal of Water Resources Development*, *32*(2), 321–338. https://doi.org/10.1080/07900627.2015.1055856

Thomas, D. E. (2005). *Developing watershed management organizations in pilot sub-basins* (Final Report). Office of Natural Resources and Environmental Policy and Planning. Ministry of Natural Resources and Environment.

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*(2), e2020WR027706. https://doi.org/10.1029/2020WR027706

Treydte, K. S., Schleser, G. H., Helle, G., Frank, D. C., Winiger, M., Haug, G. H., & Esper, J. (2006). The twentieth century was the wettest period in northern Pakistan over the past millennium. *Nature*, *440*(7088), 1179–1182. https://doi.org/10.1038/nature04743

Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and Computing*, *4*(2). https://doi.org/10.1007/BF00175354

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80–83. https://doi.org/10.2307/3001968

Wise, E. K. (2021). Sub-Seasonal Tree-Ring Reconstructions for More Comprehensive Climate Records in U.S. West Coast Watersheds. *Geophysical Research Letters*, *48*(2), e2020GL091598. https://doi.org/10.1029/2020GL091598

Woodhouse, C. A., Gray, S. T., & Meko, D. M. (2006). Updated streamflow reconstructions for the Upper Colorado River Basin. *Water Resources Research*, *42*(5), W05415. https://doi.org/10.1029/2005WR004455

World Bank. (2011). *Thailand environment monitor: Integrated water resources management - A way forward* (Technical Report No. 63368). World Bank Group.

Xu, C., Buckley, B. M., Promchote, P., Wang, S. Y. S., Pumijumnong, N., An, W., et al. (2019). Increased variability of Thailand's Chao Phraya River peak season flow and its association with ENSO variability: Evidence from tree ring $\delta^{18}O$. *Geophysical Research Letters*, *46*(9), 4863–4872. https://doi.org/10.1029/2018GL081458

Xu, C., Pumijumnong, N., Nakatsuka, T., Sano, M., & Guo, Z. (2018). Inter-annual and multi-decadal variability of monsoon season rainfall in central Thailand during the period 1804–1999, as inferred from tree ring oxygen isotopes. *International Journal of Climatology*, *38*(15), 5766–5776. https://doi.org/10.1002/joc.5859

Xu, C., Pumijumnong, N., Nakatsuka, T., Sano, M., & Li, Z. (2015). A tree-ring cellulose $\delta^{18}O$-based July–October precipitation reconstruction since AD 1828, northwest Thailand. *Journal of Hydrology*, *529*(P2), 433–441. https://doi.org/10.1016/j.jhydrol.2015.02.037

Xu, C., Sano, M., & Nakatsuka, T. (2011). Tree ring cellulose $\delta^{18}O$ of *Fokienia hodginsii* in northern Laos: A promising proxy to reconstruct ENSO? *Journal of Geophysical Research: Atmospheres*, *116*(D24). https://doi.org/10.1029/2011JD016694

Xu, C., Sano, M., & Nakatsuka, T. (2013). A 400-year record of hydroclimate variability and local ENSO history in northern Southeast Asia inferred from tree-ring $\delta^{18}O$. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *386*, 588–598. https://doi.org/10.1016/j.palaeo.2013.06.025

Xu, G., Liu, X., Trouet, V., Treydte, K. S., Wu, G., Chen, T., et al. (2019). Regional drought shifts (1710–2010) in East Central Asia and linkages with atmospheric circulation recorded in tree-ring $\delta^{18}O$. *Climate Dynamics*, *52*, 713–727. https://doi.org/10.1007/s00382-018-4215-2

Zhu, M., Stott, L. D., Buckley, B. M., Yoshimura, K., & Ra, K. (2012). Indo-Pacific Warm Pool convection and ENSO since 1867 derived from Cambodian pine tree cellulose oxygen isotopes. *Journal of Geophysical Research: Atmospheres*, *117*(D11). https://doi.org/10.1029/2011JD017198

References From the Supporting Information

Dang, T. D., Vu, D. T., Chowdhury, A. F. M. K., & Galelli, S. (2020). A software package for the representation and optimization of water reservoir operations in the VIC hydrologic model. *Environmental Modelling & Software*, *126*, 104673. https://doi.org/10.1016/j.envsoft.2020.104673

Gudmundsson, L. (2016). *Qmap: Statistical transformations for post-processing climate model output. R package version 1.0-4*.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., & Engen-Skaugen, T. (2012). Technical note: Downscaling RCM precipitation to the station scale using statistical transformations – A comparison of methods. *Hydrology and Earth System Sciences*, *16*(9), 3383–3390. https://doi.org/10.5194/hess-16-3383-2012

Harley, G. L., Maxwell, J. T., Larson, E., Grissino-Mayer, H. D., Henderson, J., & Huffman, J. (2017). Suwannee River flow variability 1550–2005 CE reconstructed from a multispecies tree-ring network. *Journal of Hydrology*, *544*, 438–451. https://doi.org/10.1016/j.jhydrol.2016.11.020

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, *7*(1), 109. https://doi.org/10.1038/s41597-020-0453-3

Hinkley, D. (1977). On quick choice of power transformation. *Applied Statistics*, *26*(1), 67. https://doi.org/10.2307/2346869

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, *99*(D7), 14415. https://doi.org/10.1029/94JD00483

Lohmann, D., Nolte-Holube, R., & Raschke, E. (1996). A large-scale horizontal routing model to be coupled to land surface parametrization schemes. *Tellus A*, *48*(5), 708–721. https://doi.org/10.1034/j.1600-0870.1996.t01-3-00009.x

Lohmann, D., Raschke, E., Nijssen, B., & Lettenmaier, D. P. (1998). Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model. *Hydrological Sciences Journal*, *43*(1), 131–141. https://doi.org/10.1080/02626669809492107

Lund, R., & Reeves, J. (2002). Detection of undocumented changepoints: A revision of the two-phase regression model. *Journal of Climate*, *15*(17), 2547–2554. https://doi.org/10.1175/1520-0442(2002)015<2547:DOUCAR>2.0.CO;2

Maxwell, R., Harley, G., Maxwell, J., Rayback, S., Pederson, N., Cook, E. R., et al. (2017). An interbasin comparison of tree-ring reconstructed streamflow in the eastern United States. *Hydrological Processes*, *31*(13), 2381–2394. https://doi.org/10.1002/hyp.11188

Robeson, S. M., Maxwell, J. T., & Ficklin, D. L. (2020). Bias correction of paleoclimatic reconstructions: A new look at 1,200+ years of Upper Colorado River flow. *Geophysical Research Letters*, *47*(1), 1–12. https://doi.org/10.1029/2019GL086689

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., et al. (2014). The NCEP climate forecast system version 2. *Journal of Climate*, *27*(6), 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., & Kitoh, A. (2012). APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society*, *93*(9), 1401–1415. https://doi.org/10.1175/BAMS-D-11-00122.1