On the Value of Paleo-Reconstructed Data: Understanding Long-Term Hydroclimatic Variability and Its Implications to Water Resources Management

Submitted by

NGUYEN Tan Thai Hung

Thesis Advisors

Assoc. Prof. Stefano GALELLI
Prof. Brendan M. BUCKLEY

Engineering Systems and Design

A thesis submitted to the Singapore University of Technology and Design in fulfillment of the requirement for the degree of Doctor of Philosophy

2020

PhD Thesis Examination Committee

TEC Chair:	Prof. Peter L. Jackson
Main Advisor:	Assoc. Prof. Stefano Galelli
Co-advisor:	Prof. Brendan M. Buckley
Internal TEC member 1:	Assoc. Prof. Bikramjit Das
Internal TEC member 2:	Asst. Prof. Ying Xu
External TEC member:	Dr. Shaowei Lin

Abstract

Engineering Systems and Design

Doctor of Philosophy

On the Value of Paleo-Reconstructed Data: Understanding Long-Term Hydroclimatic Variability and Its Implications to Water Resources Management

by NGUYEN Tan Thai Hung

Trees often grow annual rings. These rings vary year to year depending on the annual climatic conditions. Thus, tree rings are proxies through which we can infer about past climate, such as changes in the water cycle. Specifically, we can use tree rings to reconstruct history of river discharge, or streamflow, hundreds of years back in time. These long time series augment the short instrumental records and can help us better manage our water resources. While this idea was conceived decades ago, challenges in reconstruction methodologies have limited the use of streamflow reconstruction in water management. Most reconstructions rely on linear models that do not account for catchment dynamics. They are also annual, while water management decisions are made at finer time steps. My research tackles these two challenges. First, I develop a new reconstruction method based on linear dynamical systems (LDS) to account for catchment dynamics. I use the LDS method to reconstruct eight centuries of streamflow history at 62 river reaches in 16 countries across the Asian monsoon region. The reconstructions reveal that streamflow in Monsoon Asia is spatially coherent, owing to common influences by the oceans. Second, I produce seasonal and monthly reconstructions, made possible by optimally combining proxies that have different seasonal sensitivities. The monthly reconstructions of the Ping and Nan Rivers are then used to stress-test the two largest reservoirs in Thailand, proving the importance of long streamflow records containing a broad variety of droughts and pluvials. In synthesis, this work advances streamflow reconstruction methodology, provides insights about long term hydroclimatic variability of Monsoon Asia, and demonstrates the practical value of streamflow reconstruction.

Publications

Publications from this Thesis

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. (2020b). Coherent stream flow variability in Monsoon Asia over the past eight centuries — links to oceanic drivers. *Water Resources Research*, *e2020WR027*. https://doi.org/10.1029/2020WR027883

Nguyen, H. T. T., Galelli, S., Xu, C., & Buckley, B. M. (2020a). Multi-Proxy, Multi-Season Streamflow Reconstruction with Mass Balance Adjustment. *Earth and Space Science Archive*. https://doi.org/10.1002/essoar.10504791.1 (*preprint; submitted to Water Resources Research*.)

Related Publications

Libisch-Lehner, C. P., Nguyen, H. T. T., Taormina, R., Nachtnebel, H. P., & Galelli, S. (2019). On the Value of ENSO State for Urban Water Supply System Operators: Opportunities, Trade-Offs, and Challenges. *Water Resources Research*, *55*(4), 2856–2875. https://doi.org/10.1029/2018WR023622

Galelli, S., Nguyen, H. T. T., Turner, S. W. D., & Buckley, B. M. (n.d.). Time to use dendrohydrological data in water resources management? *(submitted to Journal of Water Resources Planning and Management)*

Chowdhury, A. F. M. K., Dang, T. D., Nguyen, H. T. T., Koh, R., & Galelli, S. (2020b). The Greater Mekong's climate-water-energy nexus: how ENSO-triggered regional droughts affect power supply and CO2 emissions. *Earth and Space Science Open Archive*. https://doi.org/10.1002/essoar.10504393.1 (*to appear in Earth's Future*.)

R Packages

`ldsr` Linear Dynamical System Reconstruction, available on CRAN at https://CRAN.R-project.org/package=ldsr

`mbr` Mass Balance Regression, available on GitHub at https://ntthung.github.com/mbr

`shy` Stochastic Hydrology, available on GitHub at https://ntthung.github.com/shy

Acknowledgements

I am indebted to my wonderful advisors, Stefano and Brendan. Stefano has always been fully supportive of my academic explorations and services. He knows just the right way to nudge me without making me feel stressed. He always finds a way to improve the work, from a different positioning of the paper to a missing comma. Yet, when I am too hard on myself, he knows how to lift me up. A dedicated, thoughtful, and understanding advisor that I am lucky to have for the last five years. And for the last two years, Brendan has mentored me into the wonderfully weird world of tree rings, given wings to my career, and set me on a path I couldn't have imagined. Thank you, Brendan and Stefano.

I thank the Thesis Committee for your guidance and support, both in completing this work and in my own growth as a student. Thanks to Peter and Ying for trusting me with responsibilities as a TA, and for your generous nominations for the TA awards. Peter has also been giving me all the supports that I could wish for in organizing student activities. Shaowei laid the foundation for my statistical knowledge, without which this thesis would not have been possible. And Bikram introduced me to extreme value theory, which gave me lots of ideas for future work.

PhD students are strange human beings: they speak a strange language, sleep at strange hours, and do strange things. I do all of that. But I have managed to keep my sanity, and that is all the credit of my amazing wife, Huong, and my children, Max and Dawn. It is not easy being the spouse of a PhD student, but Huong has done a miraculous job. And I thank my children for constantly asking me to leave my desk, thus bringing me back to reality.

I am grateful to my parents for raising me to be a man I am. I thank my parents and my parents-in-law for helping me to take care of the children. I have enjoyed the cheerful company of my grad school buddies, Jia Yi, Gulten, and Zhangsheng. We have done homework together, played board game together, listened to one another's talks, and had many meals together. It was through lunching and talking with you guys that I got through my first paper rejection. This thesis would not have completed without the warm support of family and friends. Thank you.

Contents

List of Figures

List of Tables

To Huong, Max, and Dawn

Prologue

As a child, I had limited access to clean water. The city's water network had not reached our neighbourhood, and we used water from a shared well. Because our groundwater was contaminated with alum, my father built our own sand and coal filter in a plastic drum to treat the water from the well. Every year I helped him cleaning the sand—an enervating task. But even after filtration, that water could only be used for washing and cleaning. The more affluent families resorted to buying bottled drinking water. My parents could not afford that, so every few days, we had to make a trip to my grandparents' place to collect tap water in a jerrycan for our most essential uses. I remember, vividly, sitting behind on dad's motorcycle, holding, or rather, hugging the container. The hassle eventually went away years later when the water mains came about.

And now, here I am, studying how to better manage water resources. I wish that in the future, no one would have to go through what I did, but I know that people in many parts of the world are suffering much worse. I want to fix that.

In searching for a better future, I became fascinated with the past—with paleohydrology. Paleohydrology is the study of the water cycle in the distant past, getting its name from the Greek words *palaios* (ancient), *hydros* (water), and *logia* (writing). The topic of paleohydrology that I am particularly interested in is streamflow reconstruction. As I will show in this thesis, understanding a catchment's past behaviour, thereby understanding its water availability, can help tremendously in planning for how we use this resource. Just as my own past has influenced my present and my future, I believe we can find the key to future water management by knowledge from catchments' history. In this thesis, I will show how we may gain such knowledge, and how we can make use of it.

Chapter 1

A Case for Streamflow Reconstruction

1.1 The Need to Extend Streamflow Records

Ninety percent of the world's population live within 10 km of an open freshwater source (lakes and rivers) (Kummu et al., 2011). Understanding the behaviour of river discharge, or *streamflow*, is crucial for managing water resources. Governmental agencies around the world have been collecting streamflow data, which are used in designing and operating critical infrastructure such as reservoirs, flood control systems and irrigation networks. Although the importance of these instrumental data is undeniable, it has been suggested that they are inadequate (Thyer et al., 2006). Streamflow is strongly influenced by climate. Climate varies over a wide range of time scales, from years to centuries or even millennia. As a result, it is expected that streamflow also varies on different range of time scales, including multi-century scales. Yet, in all countries, streamflow records are only available for several decades to about 100 years at best (Figure 1.1). Therefore, it is unlikely that instrumental records can capture all aspects of streamflow variability. For example, instrumental record may not capture extreme events that have long return periods.

These limitations have brought forth interests in the hydrological sciences community to extend the instrumental records. Trees turned out to be a reliable data source. Pioneering works of Douglass (1914, 1919) showed that trees hold records of past climate: the width of their annual rings is strongly correlated with precipitation and sun spot cycles. Strong correlations were also found between tree rings and streamflow (Hardman & Reil, 1936; Schulman, 1945, 1951), eventually leading to the first reconstructions of streamflow at the Bright Angel Creek and San Francisco River in the United States (Stockton, 1971). These reconstructions proved that it is indeed feasible to extend the instrumental streamflow record back in time using tree rings. Soon, many reconstructions were developed for the Colorado River Basin (Stockton & Jacoby, 1976), revealing that between 1900–1925, the river experienced the wettest period in the past four centuries. This abnormally pluvial period was, unfortunately, the basis for the Colorado River Compact, a water rights agreement between the southwestern U.S. states. As a result, "*water that was not likely to be in the river on a consistent basis was divided among the basin states*" (Woodhouse et al., 2006). This overallocation of water rights continues to cause severe water stresses and conflicts in the region till this day. Because of overallocation, and hence overuse, the Colorado River no longer reaches the Pacific Ocean.

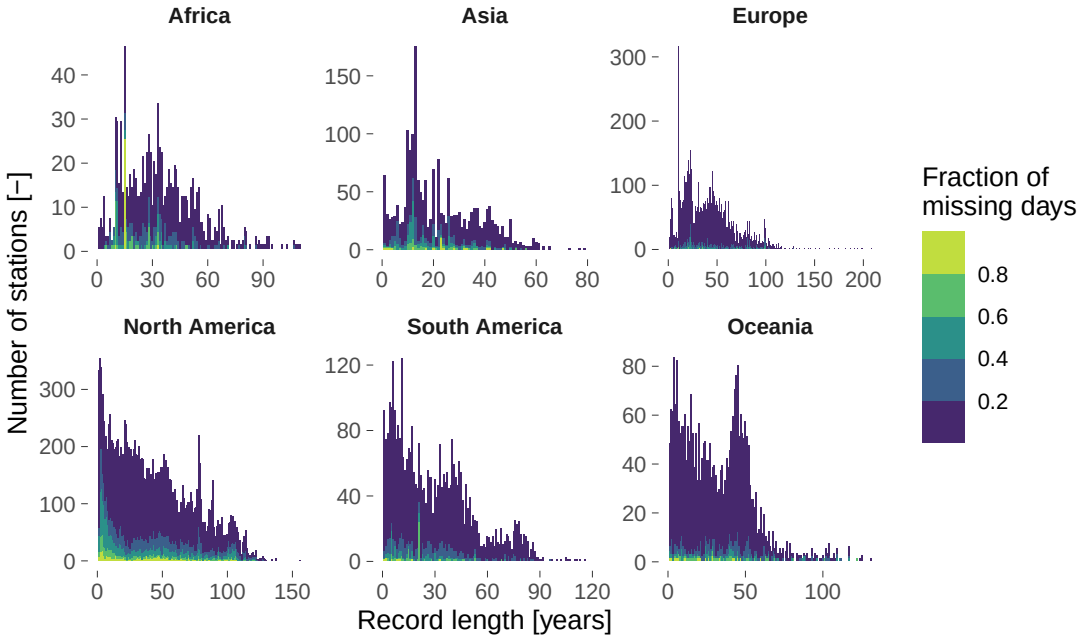

Figure 1.1. Global availability of instrumental streamflow records, analysed using the Global Streamflow Indices and Metadata Archive (Do et al., 2018).

Since Stockton and Jacoby (1976), streamflow reconstruction has brought forth insights that were unattainable with short instrumental records. Most notably, streamflow reconstructions have revealed extreme events (droughts and pluvials) in the distant past, and put recent extreme events into perspective (Meko & Woodhouse, 2011). In some cases, the paleo period was found to have more extreme droughts (e.g., Güner et al., 2017), and both more extreme droughts and pluvials (e.g., DeRose et al., 2015; Schook et al., 2016), than the instrumental period. In other cases, the opposite was observed (Woodhouse et al., 2006; Littell et al., 2016). Although varying in details, these studies—and many others (e.g., R. S. Maxwell et al., 2011; Bekker et al., 2014; Razavi et al., 2015)—came to the consensus that reconstructed streamflow data provide more understanding about streamflow variability than do instrumental data alone. Such added understanding is being transformed into water management practice in the U.S. (Meko & Woodhouse, 2011) and Canada (Sauchyn et al., 2015). Similar progress may be expected in other countries.

1.2 Physical Basis for Streamflow Reconstruction

Streamflow reconstruction relies on the science of dendrochronology—from the Greek words *dendron* (tree) and *chronos* (time). Dendrochronology studies climatic and environmental signals in tree rings, and uses those signals to date each annual ring to a calendar year. The most commonly used tree ring feature is ring width, but other features such as wood density and oxygen isotope ratio are also valuable.

1.2.1 How Do Trees Form Rings?

Trees grow under favourable climatic and environmental conditions: sufficient soil moisture and nutrients, appropriate temperature, among others. These favourable conditions often come in cycles called the growth season.

As a tree grows, it forms a tissue called *xylem*, which consists of new wood cells in the outer part of the trunk, just under the bark (xylem is derived from the Greek word *xylon*, meaning wood.) The xylem transports water and minerals from the root to the leaves. At the beginning of the growth season, the xylem is very porous so that it can transport as much water as possible. Towards the end of the growth season, however, porous wood can be dangerous: there is not enough water in the soil to fill all the pores, and cavitation can occur that ruptures the wood. As a defence mechanism, the tree increases its production of a polymer called *lignin* to thicken the cell wall so as to strengthen the wood and reduce porosity. This thickened wood, called *latewood*, is darker in colour than the wood produced earlier in the season (called *earlywood*). In the next growth season, another layer of light-coloured earlywood is formed. The colour contrast between latewood and subsequent earlywood creates a ring boundary; a ring consists of earlywood and latewood of the same growth season (Figure 1.2).

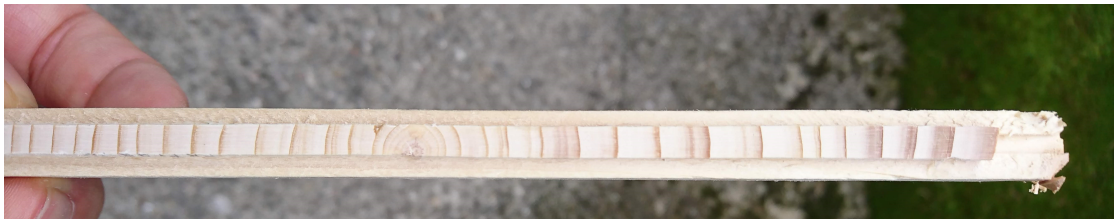

Figure 1.2. A core sample of European oak (*Quercus robur*) with annual rings and pith clearly visible. Taken in September 2018 at the 29th European Dendroecological Field Week.

1.2.2 Tree Rings and Streamflow: A Conceptual Model

Tree rings are formed with the seasonal cycles of climate. In temperate climates, temperature is often the limiting factor of tree growth. There, growth season often coincides with summer when the weather is warm enough. The warmer the temperature, the faster the growth, and the wider the rings. In tropical and sub-tropical climates, soil moisture becomes the limiting factor. Here, growth season often coincides with the rainy season when there is sustained rain to keep the soil moist. The more rain there is, the faster trees grow and the wider their rings become.

Using tree rings, we can infer about the climate and hydrology of the past. The main physical processes that involve climate and tree growth are depicted in Figure 1.3a. The climate at a given location can be characterized by precipitation and temperature, among others. These climatic inputs control soil moisture on land. Except for losses (such as groundwater recharge, evaporation, and surface runoff), the net soil moisture storage then follows two main paths: one goes out of the catchment as streamflow, the other is taken up by the trees and transpired back into the atmosphere, influencing tree growth along the way. We can thus infer a relationship between tree rings

and streamflow because they are both influenced by climate and soil moisture. Reconstructing streamflow directly from tree rings is the most common approach that has been around since Stockton (1971).

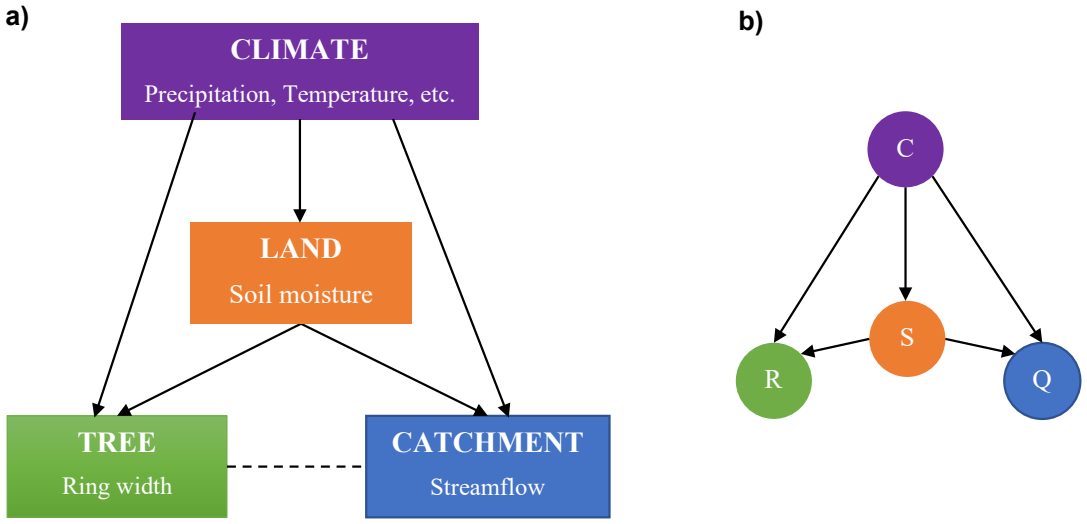

Figure 1.3. a) Relationships between hydroclimatic variables and tree growth. b) A probabilistic graphical model representing the relationships in a), where C is a vector of climate variables, S the soil moisture, R the ring width index, and Q streamflow. The arrows represent the conditional dependence among variables.

The conceptual model in Figure 1.3 also suggests two other (indirect) approaches to reconstructing streamflow: (i) from precipitation, and (ii) from soil moisture. These novel approaches have been explored recently. In (i), one first reconstructs precipitation from tree rings, then routs the precipitation reconstruction through a water balance model to obtain streamflow (Saito et al., 2015; Tozer et al., 2018; Meko et al., 2020). This approach requires a large amount of data for the water balance model's calibration. This is a key limitation to its applicability in data-scarce region such as many parts of Asia, and will not be considered further in my thesis.

Following (ii), one first needs a long term soil moisture record that is reconstructed from tree rings. Several such records have been developed (e.g., Cook et al. 1999, 2010a). They are gridded data sets of the Palmer Drought Severity Index (PDSI, W. C. Palmer, 1965; van der Schrier et al., 2013), and thus are called drought atlases. Particularly, drought atlases of North America (Cook et al., 1999, 2010b) and the Asian Monsoon region (E. R. Cook et al., 2010a) have been shown to be practical paleoclimate proxies to reconstruct streamflow. Earlier experiments used individual grid points, either in combination with ring widths (Coulthard et al., 2016) or on their own (Graham & Hughes, 2007; Adams et al., 2015). Ho et al. (2016, 2017) and Nguyen and Galelli (2018) then formalized the methodology and provided theoretical considerations. They reasoned that since both streamflow and PDSI can be modeled as functions of ring width, one can also build a model to relate streamflow to PDSI. Reconstructing streamflow from drought atlases has some important computational advantages but also has some limitations. I will discuss them in Chapter 3.

The physical discussion above yields three types of paleoclimate reconstruction: streamflow from tree rings, soil moisture from streamflow, and streamflow from soil moisture. We now derive mathematically the relationships between these reconstruction types.

Each reconstruction is a conditional distribution of one variable (e.g. streamflow) given that we have observed another variable (e.g. soil moisture), and given the historical climate. We represent these conditional distributions with a probabilistic graphical model (Koller & Friedman, 2009) as shown in Figure 1.3b. There are four random variables involved: climate (C), soil moisture (S), ring width (R), and streamflow (Q). Each of these variables can be multivariate, i.e., C includes precipitation and temperature, among others, and all variables can include multiple sites or grid points. As a convention, let $f_X(x)$ be the probability density function (PDF) of the random variable X, $f_{XY}(x,y)$ be the joint PDF of X and Y, and $f_{X|Y}(x|y)$ be the conditional PDF of X given that $Y = y$.

Reconstructing streamflow from tree ring is essentially deriving the distribution of Q given R and C, i.e, $f_{Q|R,C}(q|r,c)$, where r is the measured ring width index, and c is the historical climate. We can decompose this distribution as follows:

$$\begin{aligned} f_{Q|R,C}(q|r,c) &= \int f_{Q,S|R,C}(q,s|r,c)\, ds \\ &= \int f_{Q|S,R,C}(q|s,r,c) f_{S|R,C}(s|r,c)\, ds. \end{aligned} \tag{1.1}$$

The first equality comes from the relationship between marginal and joint distributions. The second equality comes from Bayes' theorem. Now, Q is independent of R given S and C (Figure 1.3b), so $f_{Q|S,R,C}(q|s,r,c) = f_{Q|S,C}(q|s,c)$. Consequently,

$$f_{Q|R,C}(q|r,c) = \int f_{Q|S,C}(q|s,c) f_{S|R,C}(s|r,c)\, ds. \tag{1.2}$$

Observe that $f_{Q|S,C}$ is the streamflow reconstruction from soil moisture, and $f_{S|R,C}$ is the soil moisture reconstruction from tree rings. Thus we have established mathematically the reasoning that tree-ring-based streamflow reconstruction is possible based on the link through soil moisture. $f_{Q|R,C}$ is the marginal distribution without observing the soil moisture. Instead of constructing $f_{Q|R,C}$, we can infer S from R, then Q from S, by constructing $f_{S|R,C}$ and $f_{Q|S,C}$.

1.3 Knowledge Gaps and Thesis Contributions

This thesis contributes to both the methodology and the applications of streamflow reconstruction, focusing on the Asian Monsoon region, particularly Southeast Asia. I identify four important research gaps. First, most reconstruction methods rely on linear regression, which does not account for catchment dynamics. Second, most reconstructions have an annual resolution, often insufficient for water resources applications. Third, few reconstructions have been done in Asia, all of which are single-sited; a regional, synthesized understanding is lacking. And fourth, despite interests from the water resources community, applications of streamflow reconstructions to water

system operations are still limited in scope and effectiveness. This thesis aims to fill these gaps. I present two advances in reconstruction methodology to address the first two gaps. Then, applying these new methods, I demonstrate how streamflow reconstructions provide a synoptic understanding of streamflow variability, and how high resolutions can be used to study water systems probabilistically.

In Chapter 2, I develop a novel streamflow reconstruction framework based on linear dynamical systems (LDS). The unique feature of this framework, compared to conventional reconstruction techniques, is the inclusion of a hidden state variable—the catchment flow regime (i.e., wet or dry). The LDS model also serves as a stochastic streamflow generator. Using a case study in the Ping River (Thailand), I show how we can gain insights about the regime shifts in a river basin.

In Chapter 3, I address a fundamental challenge in streamflow reconstruction: tree ring data are annual, so how do we obtain sub-annual (e.g., seasonal or monthly) reconstructions that are more practically useful for water resources management? Leveraging the fact that different tree ring proxies are sensitive to different seasons, I develop a novel framework that optimizes proxy combination in order to obtain seasonal reconstructions together with the annual one. Importantly, the framework ensures that the total seasonal flows match the annual flow closely. This mass balance criterion is important for water resources decisions such as water allocation.

In Chapter 4, I apply the LDS model to reconstruct streamflow at 62 stations on 41 rivers in 16 countries across the Asian Monsoon region—the *Monsoon Asia*. For such a large scale reconstruction, it is impractical to manually tune the model hyperparameters (the thresholds to include paleoclimate proxies), as presented in Chapter 2. Therefore, I also develop a novel, automatic, and climate-informed input selection scheme. This large scale reconstruction reveal that streamflow in Monsoon Asia is spatially coherent, owing to common influences exerted by the oceans. The reconstruction also shows how these oceanic teleconnections vary over space and time.

In Chapter 5, I first expand the method presented in Chapter 3 to obtain monthly reconstructions—a six-fold increase in temporal resolution. Then, I use the reconstructions to train a stochastic streamflow generator in order to simulate and probabilistically assess the performance of Thailand's main water resources system. The simulations produce difference performance distributions with different stochastic ensembles, suggesting that using instrumental data alone may not be adequate. This work demonstrates that, thanks to monthly reconstruction, it is now feasible to use reconstructed streamflow in operational studies of water system. The results also point out important challenges that remain to be solved so that streamflow reconstructions can be used for optimizing water system operations.

Finally, in Chapter 6, I synthesize the value of streamflow reconstruction to water resources management and discuss several interesting research direction, some of which I will pursue in my post-doctoral work.

Chapter 2

A Linear Dynamical Systems Approach to Streamflow Reconstruction Reveals History of Regime Shifts in Northern Thailand

Publication

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

2.1 Introduction

Typically, streamflow reconstruction involves two steps: one first establishes an empirical relationship between climatic proxies (e.g., tree-rings) and streamflow for the instrumental period, then feeds the paleo period's climatic proxies into the established relationship to obtain past streamflow (e.g., Hidalgo et al., 2000; Woodhouse et al., 2006; Littell et al., 2016; Allen et al., 2017). The following principal component linear regression (henceforth simply referred to as linear regression) is often adopted:

$$y_t = \alpha + \beta u_t + \varepsilon_t, \tag{2.1}$$

where t is the time index, y the streamflow (transformed to a Gaussian distribution), u the climatic proxies (typically pre-processed with Principal Component Analysis (Jolliffe, 2002), or other statistical techniques), ε a white noise, α the intercept term, and β the vector of regression coefficients. This approach has been proven to provide reliable reconstructions for a variety of modelling conditions—e.g., catchment size, length of instrumental and paleo period, hydrological regime—but it assumes that the climatic input u_t is the sole predictor of streamflow y_t. In other words, Equation 2.1 neglects the catchment dynamics and their effect on streamflow generation. As a consequence, linear regression models may not fully capture some important phenomena, such as long-range dependence in streamflow, complex flood generation mechanisms, or temporal clustering of extreme events (Pelletier & Turcotte, 1997; Koutsoyiannis, 2011). This might translate into an underestimation (or overestimation) of streamflow.

The most natural way to incorporate catchment dynamics into streamflow reconstruction is to adopt a mechanistic modelling approach. This idea was explored by several studies. Saito et al. (2015) used the Thornwaite water balance model and reconstructed seasonal temperature and precipitation records to reconstruct streamflow. Gangopadhyay et al. (2015) introduced a hybrid paleo-water balance approach consisting of two main steps: first, precipitation and temperature data are resampled to create their nonparametric reconstructions (Lall & Sharma, 1996); then, the reconstructions are fed into a water balance model to reconstruct streamflow. Tozer et al. (2018) reconstructed streamflow using a Budyko model with reconstructed data and simulated potential evaporation (PE); the lack of reconstructed PE here was compensated by using its annual average. Naturally, mechanistic models require a large amount of hydrological data, either instrumental, simulated, or reconstructed. Such data may not always be available with the required spatial and temporal resolution. This limits the applicability of mechanistic models.

Recently, Bracken et al. (2016) developed a hidden Markov model for streamflow reconstruction. The hidden state is derived from climate proxies and interpreted as the "state of the climate"; streamflow is then reconstructed from the climate state via a log-linear function. In this hierarchical model, streamflow generation depends on climate dynamics rather than catchment dynamics.

The main motivation for this work is to develop a streamflow reconstruction technique that accounts explicitly for the catchment dynamics without requiring a substantial amount of data. We address this challenge by appealing to linear dynamical systems—a class of models that has been used widely in hydrology, as it provides a good approximation for many natural phenomena, including the rainfall-runoff process (e.g., Cooper & Wood, 1982a; Ramos et al., 1995). Specifically, we model the relationship between climatic proxies and streamflow using the state-space representation of a discrete, linear dynamical system, which allows us to account for the dynamics of the catchment state as well as the effect of both climate proxies and catchment state on the streamflow generation process. Traditionally, linear dynamical systems are learned using the Expectation-Maximization (EM) algorithm (e.g., Cheng & Sabes, 2006; Shumway & Stoffer, 1982). However, EM cannot be used directly for streamflow reconstructions because the length of the climate proxies differs from that of the streamflow time series. To overcome this, we propose a novel variant for the EM algorithm that accounts for missing data. In this variant, model learning and reconstruction are carried out simultaneously instead of the two-step approach described in the first paragraph.

The technique is tested in the Ping River Basin (northern Thailand), where we reconstruct 406 years of annual streamflow based on the time series of the Palmer Drought Severity Index, retrieved from the Monsoon Asia Drought Atlas (E. R. Cook et al., 2010a). We show that the proposed model yields two important advantages. First, the reconstructed streamflow time series is complemented by a corresponding time series of a catchment state variable that provides information on the catchment's dynamics (e.g., regime-like behaviour), thereby assisting with the analysis of historical droughts and pluvials. Second, the model can readily generate stochastic streamflow

replicates. These advantages are obtained with similar skills and a marginal increase in computational requirements compared to linear regression.

2.2 Study Site and Data

2.2.1 The Ping River Basin

The Ping River, located in northern Thailand, drains a catchment of 33,900 km^2 (Komori et al., 2012). Along with the Nan River, the Ping is one of the main tributaries of the Chao Phraya, whose basin covers 30% of the country's surface (Figure 2.1a). The water flowing in the Chao Phraya Basin serves multiple users—i.e., agricultural, industrial, and domestic supply, hydropower generation, navigation, and prevention of seawater intrusion in the Gulf of Thailand—supporting a population of approximately 25 million people, including 8 million in Bangkok (Divakar et al., 2011; Takeda et al., 2016). A key component of the water system is the Bhumibol Reservoir, located on the Ping River. The reservoir has a large active storage capacity—about 9,700 Mm^3—that helps control floods and meet the demand of the different water users.

In this study, we aim to reconstruct annual streamflow (on a water year basis, April to March) at the P.1 stream gauging station, located in Chiang Mai, upstream of Bhumibol Reservoir (Figure 2.1a). In this area, monthly streamflow exhibits a strong seasonal pattern, with higher flow observed during the Southwest Monsoon season (early May to October-November) (Figure 2.1b). Peak flows and, therefore, floods are generally observed during the second part of the monsoon season when heavy rainfall events occur over a catchment that is already wet. The flood generation mechanism can vary on an annual basis, as it depends on the intertwining interactions between monsoon rainfall, global circulations, and tropical storms (Lim & Boochabun, 2012). For instance, the 2006 flood appears to be caused solely by monsoon rainfall—whose intensity is amplified in La Niña years (Kripalani & Kulkarni, 1997)—while larger events, such as those that occurred in 1973 and 2005, were caused by the combination of monsoon rainfall and tropical storm activities (Lim & Boochabun, 2012). Naturally, such complex streamflow generation process makes the reconstruction exercise difficult, especially when using data derived from moisture-limited trees, because saturated overland flow is not reflected in the tree-ring widths.

We obtained monthly streamflow data at P.1 from the Thai Royal Irrigation Department's database (http://www.hydro-1.net). To match the last year of our paleo data source (described in the next section), we used 85 water years—from April 1921 to March 2005. Since water management in Thailand is carried out on a water year basis, we believe that a reconstruction based on a water year is more useful. We filled in the single missing data point (June 1933) by averaging the monthly streamflow of the previous and subsequent month.

2.2.2 The Monsoon Asia Drought Atlas

In Southeast Asia, streamflow reconstruction studies are rare because the necessary instrumental data for calibration are often limited and, most importantly, tree-ring records are scarce—an issue partially attributable to the lack of suitable tree species

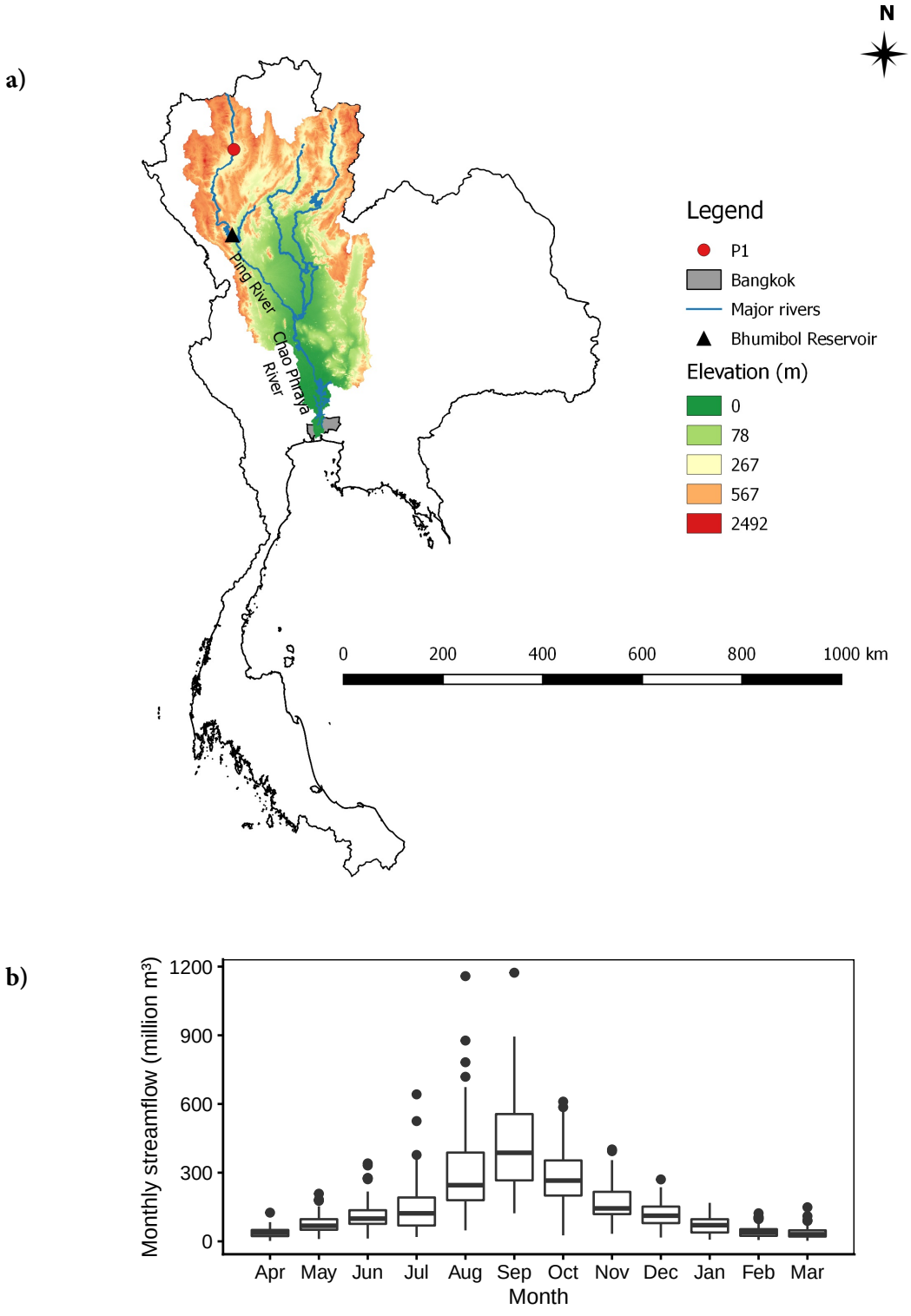

Figure 2.1. a) Map of the Chao Phraya River Basin and main tributaries, including the Ping River. The P.1 station is indicated with a red dot. b) Distribution of the monthly streamflow measured at P.1 for the period April 1921 to March 2005.

(Sano et al., 2009). In fact, to the authors' knowledge, there has been only one streamflow reconstruction attempt in Southeast Asia (D'Arrigo et al., 2011a). To address the problem of data scarcity, we propose to use the Palmer's Drought Severity Index (PDSI). While there are only a few tree-ring sites in Southeast Asia, the PDSI is available in a gridded data set called the Monsoon Asia Drought Atlas (MADA) (E. R. Cook et al., 2010a)—a spatial-temporal drought map over the Asian Monsoon region, with resolution $2.5° \times 2.5°$. The map comprises 534 grid points, each containing an annual time series of the PDSI, from 1300 to 2005, reconstructed from tree-ring chronologies. The theoretical ground for using the PDSI as a climate proxy (recall Section 1.2.2) is that both PDSI and streamflow can be expressed as regression functions of tree-rings; hence, one can build a regression function between them. Based on this idea, Ho et al. (2016) utilized the Living Blended Drought Atlas (LBDA) (E. R. Cook et al., 2010b)—a grid of PDSI time series reconstructed from tree-rings over North America—to reconstruct streamflow in the Missouri River Basin, yielding good reconstruction skill (adjusted R^2 ranged between 0.56 and 0.90).

Preliminary analyses showed that annual streamflow at P.1 station has higher correlation with the nearby MADA grid points than with nearby tree-ring chronologies (Table 2.1, Figure 2.2). The analyses also showed that 1,200 km is the optimal search radius to include the MADA grid points (Figure 2.3). There are three possible explanations for this result. First, this radius incorporates valuable information from the Bidoup-Nui Ba tree-ring site in Vietnam, about 1,200 km south-east of P.1. The chronology from this site was a major "anchor" for PDSI reconstruction in this region (E. R. Cook et al., 2010a). Second, the chronologies at the southern-most end of the Tibetan Plateau may have contributed to the reconstruction (Figure 2.2) [*Cook, 2017, personal communication*]. Finally, going beyond 1,200 km means leaving the climate zone characterizing the region (Peel et al., 2007). Based on these analyses, we used 51 MADA grid points (within the optimal radius) for the period 1600–2005 as the paleoclimatic data for our streamflow reconstruction. The use of a shorter time series is justified by the fact that most tree-ring chronologies in Southeast Asia started from the 17^{th} century onwards (Buckley et al., 2007b; Sano et al., 2009)—so, data for the period before 1600 may be less reliable.

2.3 Linear Dynamical System Learning–Reconstruction

In this section, we first provide a brief overview of linear dynamical systems, and then describe in greater detail our proposed variant of the Expectation-Maximization algorithm used for the reconstruction exercise. We then show how the linear dynamic model can be used to generate stochastic streamflow replicates. Finally, we describe the implementation of the algorithm in the R package `ldsr`.

Table 2.1. Correlation between proxies (MADA grid points and tree-ring chronologies[a]) and P.1 streamflow, arranged by increasing distance from station P.1, grouped by similar distances. The numerical IDs (e.g., 252) are those of the MADA grid points. The other IDs belong to the tree ring chronologies.

ID	Starting year	Ending year	Distance to P.1 (km)	Correlation	p-value
252	1300	2005	27	0.52	< 1E-4
TH001	1558	2005	55	0.2	0.0632
TH006	1648	2004	85	-0.04	0.7464
251	1300	2005	282	0.53	< 1E-4
TH002	1786	1993	354	0.13	0.2757
275	1300	2005	368	0.48	< 1E-4
TH003	1616	1993	370	-0.04	0.7589
LS001	1743	2005	407	-0.22	0.0407
TH004	1693	2006	423	0.18	0.0919
LS002	1785	2005	439	-0.14	0.1863
301	1300	2005	500	0.42	0.0001

[a] Standardized chronology indices are obtained from the dendrobox project (dendrobox.org; Zang, 2015)

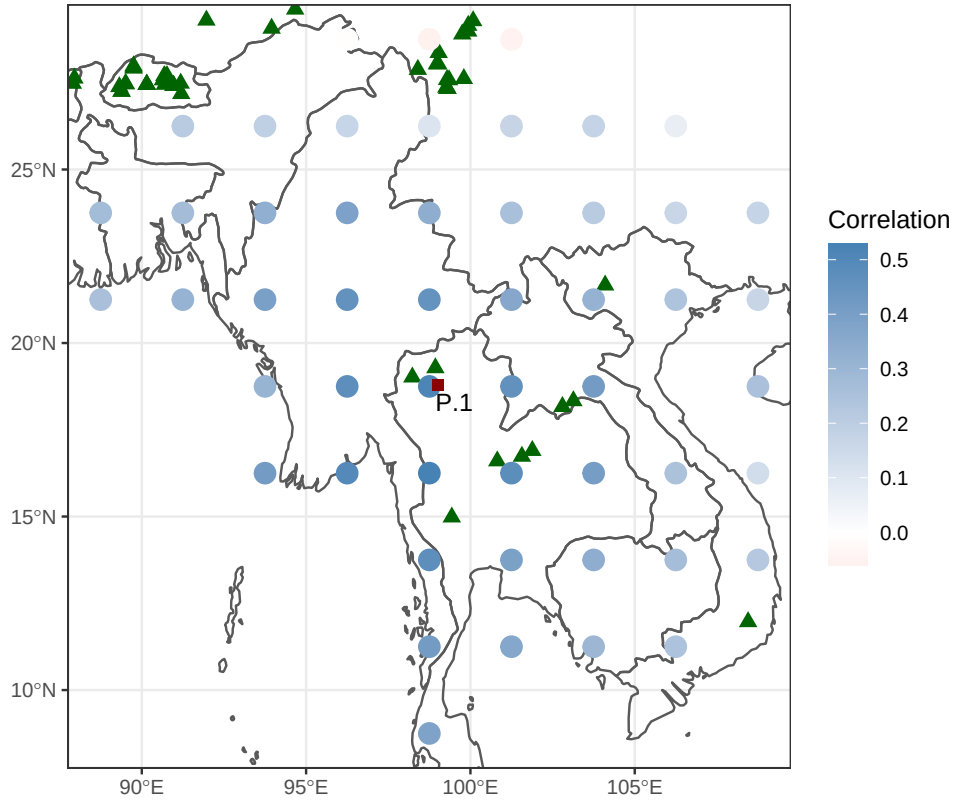

Figure 2.2. Map showing MADA grid points (colour-scaled circles) within 1,200 km of station P.1 (red square), and nearby tree-ring sites (green triangles). The MADA grid points show a radially decreasing correlation pattern. Beyond 1,200 km, correlation is insignificant. (Annual correlation between streamflow and PDSI was calculated for the instrumental period 1921–2005.)

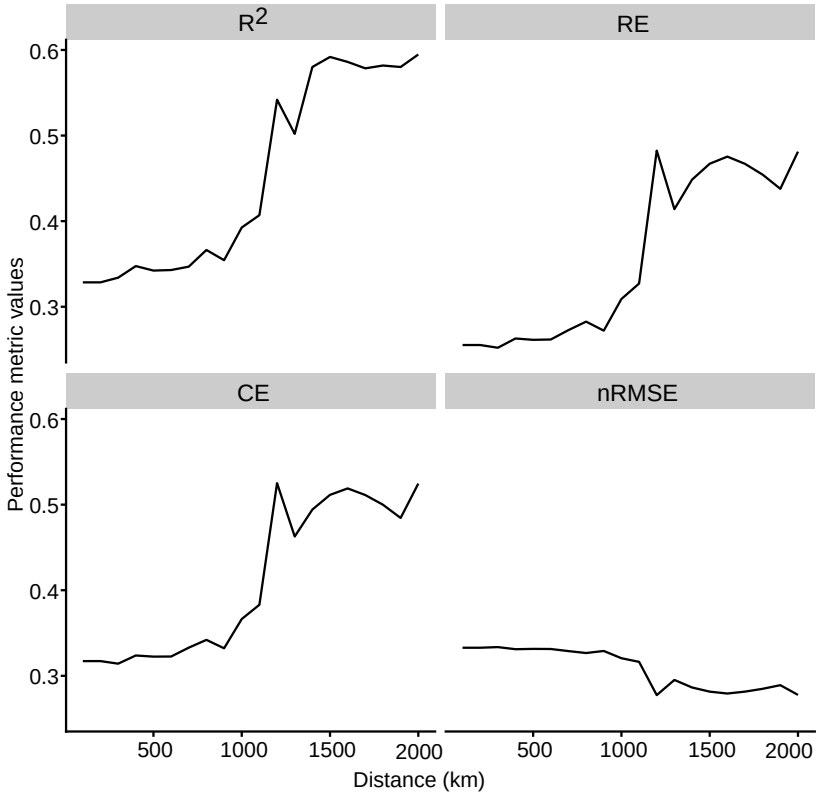

Figure 2.3. Results of the preliminary analysis to find the optimal search radius. For each radius between 100–2000 km with 100 km increments, we fitted the benchmark model and calculated the performance scores (Section 2.3.5). This figure shows the changes in each score with increasing radius.

2.3.1 Linear Dynamical Systems

We consider a stochastic, discrete, time-invariant, linear dynamical system with the following state-space representation:

$$x_{t+1} = Ax_t + Bu_t + w_t \tag{2.2}$$
$$y_t = Cx_t + Du_t + v_t \tag{2.3}$$
$$w_t \sim \mathcal{N}(0, Q)$$
$$v_t \sim \mathcal{N}(0, R)$$

where $x \in \mathbb{R}^p$ is the system state; $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^q$ are the input and output; $w \in \mathbb{R}^p$ and $v \in \mathbb{R}^q$ are both white noise, independent of each other. Henceforth, we refer to the system governed by Equations 2.2 and 2.3 as a linear dynamical system (LDS), and its parameters $A \in \mathbb{R}^{p\times p}$, $B \in \mathbb{R}^{p\times m}$, $C \in \mathbb{R}^{q\times p}$, $D \in \mathbb{R}^{q\times m}$, $Q \in \mathbb{R}^{p\times p}$, and $R \in \mathbb{R}^q$ are collectively referred to as θ. Furthermore, we assume that, at time $t = 1$, the system starts from an initial state $x_1 \sim \mathcal{N}(\mu_1, V_1)$. Note that the LDS model is a state-space representation of the ARMAX model (Auto-Regressive–Moving Average with eXogeneous

input) (Ramos et al., 1995; Shumway & Stoffer, 2011), but it has an advantage over ARMAX: the system state is modeled explicitly. In rainfall-runoff modeling applications (e.g., Ramos et al., 1995), x, u, and y represent the catchment state (an indicator of its wetness/dryness), rainfall, and runoff (or streamflow), respectively. In the context of this study, x and y maintain the same meaning, whereas the input u is represented by the climatic proxy.

The model can be used for both single- and multi-site applications (Cooper & Wood, 1982a, 1982b). In the latter case, $x_t^{(j)}$ and $y_t^{(j)}$ represents the state and output at the j-th site, and the matrices A, C, Q and R capture the spatial dependence between the sites. Moreover, Equation 2.2 implies that state transition is a first-order Markov process. Higher order Markov processes can be transformed into first-order by expanding the state space. For example, in this case study, a one-dimensional system state x_t represents the catchment wetness at time t. We may use a two-dimensional system state to represent the catchment wetness at the current and previous time step. The dimensions of matrices A, B, C, D, Q, and R must then be increased accordingly.

One observes that linear regression is a sub-class of the LDS model: the state-dependent term Cx in Equation 2.3 is replaced by the constant intercept term α in Equation 2.1, and the state transition Equation 2.2 is unused in linear regression. As a result, linear regression may not fully capture phenomena related to the catchment dynamics, such as flood generation mechanisms or long-range dependence (Koutsoyiannis, 2011). In this respect, LDS is better suited, since it uses the information regarding both catchment state and climate proxies to estimate streamflow. Another key advantage of LDS over linear regression concerns the autocorrelation structure of the output variables. When the linear model is learned using least square estimators, as is often the case, serial independence is implicitly assumed (DeGroot & Schervish, 2012, pg. 701), which is often not valid for climatic and hydrological processes (see Pelletier & Turcotte, 1997). This, on the other hand, is not a problem for the LDS model, which is learned using a maximum likelihood method, as we shall see in Section 2.3.2.

2.3.2 Learning the System States and Parameters with the Expectation Maximization Algorithm

Equations 2.2 and 2.3 indicate that in order to model the annual streamflow (i.e., the system output y), input u, state x, and parameters θ must be known. When the state trajectory is known, the task of estimating the parameters is generally referred to as the *system identification* problem (Roweis & Ghahramani, 2001). When the system parameters are known, the task of estimating the state trajectory is called the *state estimation* problem [*ibidem*]. Interestingly, the task at hand is a combination of both: only the system output y and input u (i.e., tree rings or soil moisture) are available, so neither the state nor the parameters are known. One possible solution, first proposed by Shumway and Stoffer (1982), is to iterate between state estimation and system identification using the Expectation-Maximization (EM) algorithm (Dempster et al., 1977). This method was further developed by Ghahramani and Hinton (1996) and Cheng and Sabes (2006). The algorithm starts with an initial parameter set $\hat{\theta}_0$. At the k^{th} iteration, given the

current parameter set $\hat{\theta}_k$, the Expectation step (E-step) estimates the hidden states

$$\hat{X}(\hat{\theta}_k) = \mathbb{E}\left[X \mid Y, \hat{\theta}_k\right];$$

where $t = 1, ..., T$ is the time index, $X = (x_1, ..., x_T)$ is the state trajectory, $Y = (y_1, ..., y_T)$ is the output trajectory, and the hat notation denotes the estimator for the respective unknown quantity. In other words, this step solves the state estimation problem. Then, given the newly estimated state, the Maximization step (M-step) finds a new parameter set $\hat{\theta}_{k+1}$ that maximizes the likelihood of the output; that means, the M-step solves the system identification problem. Mathematically, the goal of the M-Step is to find

$$\hat{\theta}_{k+1} = \arg\max \mathcal{L}\left(Y, X \mid \hat{\theta}_k, U\right)$$

where $\mathcal{L}(.)$ denotes the likelihood function, and $U = (u_1, ..., u_T)$ the input trajectory. The critical property of the EM formulation is that the likelihood is non-decreasing after each iteration step (Dempster et al., 1977), so Equation 2.4 always holds.

$$\mathcal{L}\left(Y, X \mid \hat{\theta}_{k+1}, U\right) - \mathcal{L}\left(Y, X \mid \hat{\theta}_k, U\right) \geq 0 \tag{2.4}$$

EM iterates between the E-step and the M-step until convergence, i.e., when the left-hand-side of Equation 2.4 is less than a predetermined threshold τ. In the case of Gaussian likelihood, convergence is always guaranteed (Wu, 1983). In the remainder of this section, we describe the mathematical details of the EM algorithm.

2.3.2.1 E-step

Throughout the E-step, the system parameters are kept at the values determined in the previous M-step. Given the observed output trajectory Y, the state trajectory X is estimated using the Kalman smoother (Anderson & Moore, 1979). Let

$$\hat{x}_{t|s} = \mathbb{E}\left[x_t \mid y_1, ..., y_s\right] \tag{2.5}$$
$$\hat{V}_{t|s} = \mathbb{V}\left[\hat{x}_{t|s} \mid y_1, ..., y_s\right] \tag{2.6}$$

Thus, $\hat{x}_{t|s}$ is the estimated state at time t given observations up to time s, and $\hat{V}_{t|s}$ is the estimated variance of that state estimator. When $s > t$, the estimation task is called *smoothing*, when $s < t$, it is called *prediction*, and when $s = t$, it is called *filtering*. The overall goal of the Kalman smoother is to compute $\hat{x}_{t|T}$ (hence the name *smoother*, as $T \geq t$). This task is done in two passes: forward and backward.

The forward pass utilizes the Kalman filter (Kalman, 1960) to estimate $\hat{x}_{t|t}$ (hence the name *filter*). First, we assume an initial state $x_1 \sim \mathcal{N}(\mu_1, V_1)$, so $\hat{x}_{1|1} = \mu_1$ and $\hat{y}_{1|1} = C\mu_1 + Du_1$. For $t = 2, ..., T$, given the latest available estimate $\hat{x}_{t-1|t-1}$ based on observations up to time $t-1$, we predict the current state using Equation 2.2:

$$\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1} + Bu_t \tag{2.7}$$
$$\hat{V}_{t|t-1} = A\hat{V}_{t-1|t-1}A' + Q \tag{2.8}$$

The system output for the current time step is predicted using Equation 2.3:

$$\hat{y}_{t|t-1} = C\hat{x}_{t|t-1} + Du_t \tag{2.9}$$

Once the actual output y_t is observed, the difference between the predicted output (Equation (2.9) and observation is useful for improving the state estimation:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(y_t - \hat{y}_{t|t-1}) \tag{2.10}$$

where

$$K_t = \hat{V}_{t|t-1}C'(C\hat{V}_{t|t-1}C' + R)^{-1} \tag{2.11}$$

is the Kalman gain. The computation in Equation 2.10 is called *measurement update*, which adds an updating term to $\hat{x}_{t|t-1}$ to obtain $\hat{x}_{t|t}$. Equation 2.10 also shows that the updating term is proportional to the prediction error

$$\delta_t := y_t - \hat{y}_{t|t-1} \tag{2.12}$$

Finally, the variance of the state estimation can be updated as well

$$\hat{V}_{t|t} = (I - K_tC)\hat{V}_{t|t-1} \tag{2.13}$$

where I is the identity matrix. One can think of the distribution of $\hat{x}_{t|t-1}$ as the prior distribution of x_t, and the distribution of $\hat{x}_{t|t}$ as the posterior distribution, once new data y_t is obtained. Furthermore, the Kalman filter can be shown to be the optimal estimator, in that it minimizes the mean squared error. The detailed proofs can be found in Shumway and Stoffer (2011, Chapter 6).

The Kalman filter is the optimal estimator for x_t given all observations up to time t. However, if one has all the observations $y_1, ..., y_T$, one can improve the state estimation further using the Rauch-Tung-Striebel (RTS) recursion (Rauch et al., 1965) in the backward pass. This pass is initialized with $\hat{x}_{t|T}$ and $\hat{V}_{t|T}$ from the forward pass. For $t = T-1, ..., 1$, the following quantities are computed

$$J_t = \hat{V}_{t|t}A(\hat{V}_{t+1|t})^{-1} \tag{2.14}$$

$$\hat{x}_{t|T} = \hat{x}_{t|t} + J_t(\hat{x}_{t+1|T} - \hat{x}_{t+1|t}) \tag{2.15}$$

$$\hat{V}_{t|T} = \hat{V}_{t|t} + J_t(\hat{V}_{t+1|T} - \hat{V}_{t+1|t})J_t' \tag{2.16}$$

$$\hat{y}_{t|T} = C\hat{x}_{t|T} + Du_t \tag{2.17}$$

In the forward pass, one updates the state estimation based on $(y_t - \hat{y}_{t|t-1})$. In the backward pass, one does so based on $(\hat{x}_{t+1|T} - \hat{x}_{t+1|t})$. The multiplier J_t acts as a *gain*, similarly to the Kalman gain in Equation 2.11.

2.3.2.2 M-step

Throughout the M-step, the state values are fixed as those obtained in the last E-step. The goal of the M-step is to find the maximum likelihood estimators for the system

parameters. If the state trajectory X were known, the expression for the log-likelihood would be

$$\begin{aligned} L = \log \mathcal{L}\left(Y, X \mid \theta, U\right) = &- \sum_{t=1}^{T} \frac{1}{2}(y_t - Cx_t - Du_t)' R^{-1}(y_t - Cx_t - Du_t) \\ &- \sum_{t=2}^{T} \frac{1}{2}(x_t - Ax_{t-1} - Bu_{t-1})' Q^{-1}(x_t - Ax_{t-1} - Bu_{t-1}) \\ &- \frac{1}{2}(x_1 - \mu_1)' V_1^{-1}(x_1 - \mu_1) - \frac{1}{2}\log|V_1| \\ &- \frac{T}{2}\log|R| - \frac{T-1}{2}\log|Q| - \frac{T(p+q)}{2}\log 2\pi \end{aligned} \tag{2.18}$$

where p and q are the dimensions of the state and output vectors x and y. However, X is not known, but its distribution $p(X \mid Y, \hat{\theta}_k)$ is available from the E-step. Therefore, we maximize the *expected* log-likelihood

$$\mathcal{Q} = \mathbb{E}[L] = \mathbb{E}\left[\log \mathcal{L}\left(Y, X \mid \theta, U\right)\right] \tag{2.19}$$

Observe that $\mathcal{Q}$ is a sum of quadratic terms, i.e., the log-likelihood is a concave function, because the relationships in Equations 2.2 and 2.3 are linear, and the noises are Gaussian. Thus, the parameters can be determined analytically by taking the derivative of the log-likelihood and setting it to 0; the solution is a global optimum. Since the analytical expressions for the updated parameters are quite cumbersome, some further shorthand notations are necessary. Let

$$\begin{aligned} \hat{x}_t &= \hat{x}_{t|T} \\ P_t &= \hat{x}_t \hat{x}_t' + \hat{V}_{t|T} \\ P_{t,s} &= \hat{x}_t \hat{x}_s' + \mathrm{Cov}(x_t, x_s) \end{aligned}$$

where s, t are time step indices and

$$\mathrm{Cov}(x_t, x_s) = \hat{V}_{t|T} J_s'$$

Lemma 1. *(Cheng & Sabes, 2006) The update equations for all parameters are*

$$[A\ B] = \left[\sum_{t=1}^{T-1} P_{t+1,t} \quad \sum_{t=1}^{T-1} \hat{x}_{t+1} u_t'\right] \begin{bmatrix} \sum_{t=1}^{T-1} P_t & \sum_{t=1}^{T-1} \hat{x}_t u_t' \\ \sum_{t=1}^{T-1} u_t \hat{x}_t' & \sum_{t=1}^{T-1} u_t u_t' \end{bmatrix}^{-1} \tag{2.20}$$

$$[C\ D] = \left[\sum_{t=1}^{T} y_t \hat{x}_t' \quad \sum_{t=1}^{T} y_t u_t'\right] \begin{bmatrix} \sum_{t=1}^{T} P_t & \sum_{t=1}^{T} \hat{x}_t u_t' \\ \sum_{t=1}^{T} u_t \hat{x}_t' & \sum_{t=1}^{T} u_t u_t' \end{bmatrix}^{-1} \tag{2.21}$$

$$Q = \frac{1}{T-1} \sum_{t=1}^{T-1} (P_{t+1} - A P_{t,t+1} - B u_t \hat{x}_{t+1}') \tag{2.22}$$

$$R = \frac{1}{T} \sum_{t=1}^{T} (y_t - C\hat{x}_t - D u_t) y_t' \tag{2.23}$$

$$\mu_1 = \hat{x}_1 \tag{2.24}$$

$$V_1 = \hat{V}_1 \tag{2.25}$$

Cheng and Sabes (2006) presented the results without proof. We provide a proof in Appendix A.1.

The EM algorithm is summarized in **Algorithm 1**. It requires the system input and output trajectory (Y and U), and returns the parameter set $\hat{\theta}$ and the estimated state and output trajectory ($\hat{X}$ and $\hat{Y}$). Note that the solution returned by the EM algorithm is a local optimum—since the global optimum found at each M-step is optimal for that M-step only, and may not correspond to the global optimum of the overall problem.

Algorithm 1 Learning a linear dynamical system with the expectation–maximization algorithm

Require: Y, U
$k = 0$
Initialize $\hat{\theta}_0$
Initialize x_1 and $\hat{y}_1$.
repeat
  for $t = 2, ..., T$ **do**
    $\hat{x}_{t|t} \leftarrow \mathbb{E}(x_t \mid y_1, ..., y_t, \hat{\theta}_k)$ ▷ Kalman filter (Equations 2.7 to 2.13)
  end for
  for $t = T-1, ..., 1$ **do**
    $\hat{x}_{t|T} \leftarrow \mathbb{E}(x_t \mid y_1, ..., y_T, \hat{\theta}_k)$ ▷ RTS recursion (Equations 2.14 to 2.17)
  end for
  $\hat{\theta}_{k+1} = \arg\max \mathcal{L}(Y, X \mid \hat{\theta}_k, U)$ ▷ M-Step (Theorem 1)
  $k = k + 1$
until $\mathcal{L}\left(Y, X \mid \hat{\theta}_{k+1}, U\right) - \mathcal{L}\left(Y, X \mid \hat{\theta}_k, U\right) \leq \tau$ ▷ Convergence
Return: $\hat{X}, \hat{Y}, \hat{\theta}$

2.3.3 Simultaneous Learning–Reconstruction

Typically, a reconstruction problem is solved in two phases: learning and reconstruction. Accordingly, the study horizon should be divided into two parts: the paleo period (with $-T_p \leq t \leq 0$) and the instrumental period ($1 \leq t \leq T$), as illustrated in Figure 2.4a. Learning involves building a regression model for the instrumental period. Reconstruction is then carried out by feeding the paleo period's input into the regression model to obtain the paleo period's output. Although this conventional approach works well for linear regression, it is not suitable for LDS models because of two issues. First, the EM algorithm not only learns the system parameters, but it also derives an estimate for the initial state x_1, which is necessary to commence the state transition. When the LDS is learned with only the instrumental period's data, the modeller faces a question in the reconstruction phase: which initial state to use at time $t = -T_p$? As it turns out, this is not a major problem. Equation 2.2 implies that the state transition is Markovian. Thus, regardless of the initial state at time $t = -T_p$, the effect of the initial conditions diminishes as the system evolves through time, and, eventually, the state trajectory converges. One therefore just needs to discard the initial transition period.

The second, and most critical, problem arises when the paleo period meets the instrumental one. At this point in time, the system state may be different from the estimated x_1 (see Figure 2.4b). While the estimated θ is optimal for the original x_1 (derived in the learning phase), it may not be optimal for the new x_1 (derived in the reconstruction phase). Worse still, if the system is propagated further into the instrumental period, the state trajectory may also be different from its original estimation in the learning phase, effectively invalidating the learned model. It is not possible to force the system to the desired x_1 because once the system parameters are given, the system is only driven by the input.

To solve this issue, we can drop the paleo/instrumental period delineation and provide the EM algorithm with the entire input time series (Figure 2.4c). Since the time spans of the input (climatic proxy) and output (instrumental data) no longer match each other, modifications to the EM algorithm are necessary to handle the missing observations. In the forward pass (i.e, the Kalman filter step), measurement update is skipped. In the M-step, we make use of Lemma 2.

Lemma 2. *In the M-step, only the observed part, $Y_{\mathcal{O}}$ (and not the missing part, $Y_{\mathcal{M}}$), is necessary to calculate the expected likelihood in Equation 2.19. The update equations for A, B, Q, x_1, and V_1 remain the same. The update equations for C, D, and R are*

$$\begin{bmatrix} C & D \end{bmatrix} = \begin{bmatrix} \sum_{t\in\mathcal{O}} y_t \hat{x}_t' & \sum_{t\in\mathcal{O}} y_t u_t' \end{bmatrix} \begin{bmatrix} \sum_{t\in\mathcal{O}} P_t & \sum_{t\in\mathcal{O}} \hat{x}_t u_t' \\ \sum_{t\in\mathcal{O}} u_t \hat{x}_t' & \sum_{t\in\mathcal{O}} u_t u_t' \end{bmatrix}^{-1} \tag{2.26}$$

$$R = \frac{1}{|\mathcal{O}|} \sum_{t\in\mathcal{O}} (y_t - C\hat{x}_t - Du_t) y_t' \tag{2.27}$$

The proof is provided in Appendix A.2.

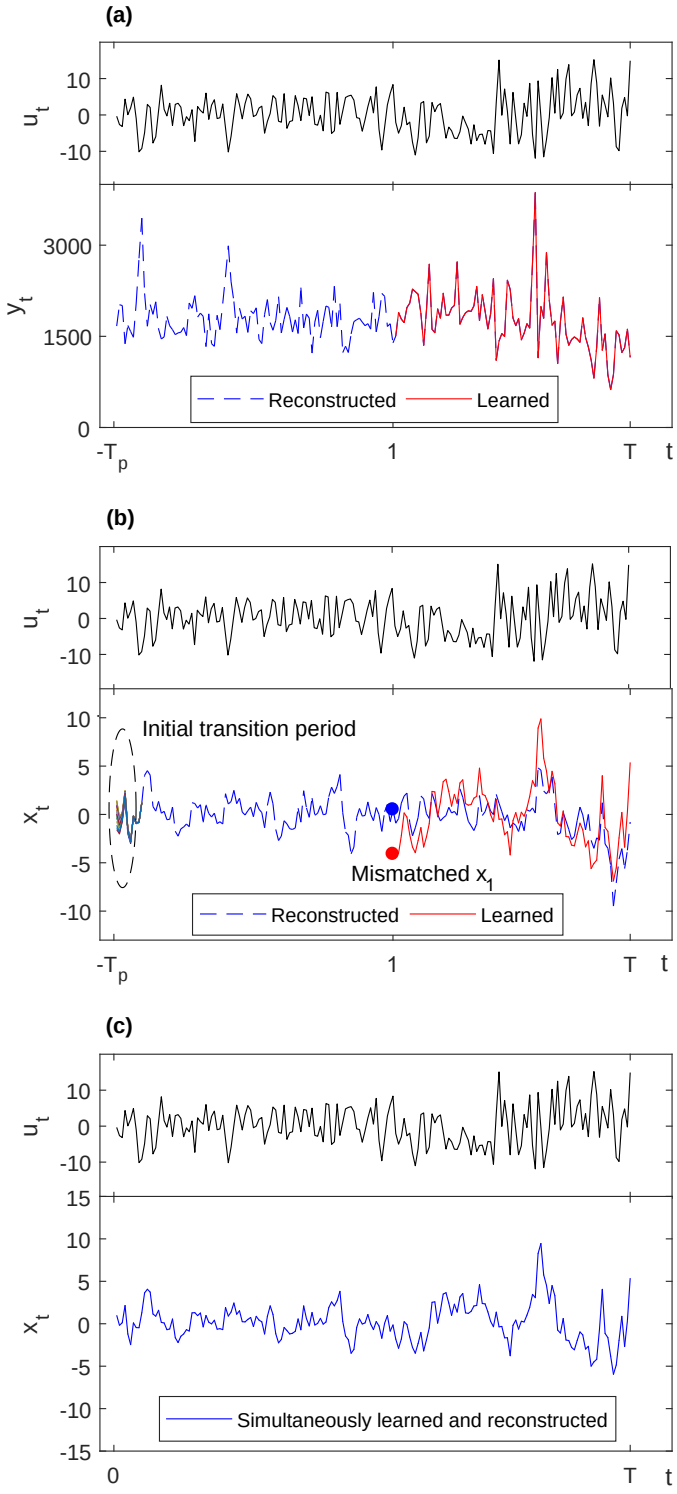

Figure 2.4. Motivation for simultaneous learning-reconstruction. a) Conventional learning-reconstruction delineation: the model is first learned using the instrumental period's data ($1 \leq t \leq T$), and then used to reconstruct streamflow in the paleo period ($-T_p \leq t \leq 0$). b) When this delineation is used for the linear dynamical systems model, two problems arise: (i) the initial transition period must be discarded, and (ii) when the state trajectory is propagated from $t = -T_p$, its value at $t = 1$ may mismatch with its estimated value during the learning phase. c) Simultaneous learning-reconstruction eliminates these problems.

With this modification, a new estimation for y in the entire study horizon is created at each iteration. As a result, when the EM algorithm converges, the system state and parameters are learned, and the reconstruction is completed at the same time. Simultaneous learning-reconstruction is achieved. The modified algorithm is summarized in Algorithm 2.

Algorithm 2 Simultaneous learning-reconstruction with the Expectation–Maximization algorithm

Require: Y, U
$k = 0$
Initialize $\hat{\theta}_0$
Initialize x_1
repeat
  for $t = 2, ..., T$ **do** ▷ Kalman filter (Equations 2.7 to 2.13)
    if $y_t \neq \text{NA}$ **then**
      $\hat{x}_{t|t} \leftarrow \mathbb{E}(x_t \mid y_1, ..., y_t, \hat{\theta}_k)$
    else
      $\hat{x}_{t|t} \leftarrow \mathbb{E}(x_t \mid y_1, ..., y_{t-1}, \hat{\theta}_k)$ ▷ (Equation 2.9)
    end if
  end for
  for $t = T - 1, ..., 1$ **do** ▷ RTS recursion (Equations 2.14 to 2.17)
    $\hat{x}_{t|T} \leftarrow \mathbb{E}(x_t \mid y_1, ..., y_T, \hat{\theta}_k)$
  end for0
  $\hat{\theta}_{k+1} = \arg\max \mathcal{L}(Y_\mathcal{O}, X \mid \hat{\theta}_k, U)$ ▷ M-Step (Theorem 2)
  $k = k + 1$
until $\mathcal{L}\left(Y_\mathcal{O}, X \mid \hat{\theta}_{k+1}, U\right) - \mathcal{L}\left(Y_\mathcal{O}, X \mid \hat{\theta}_k, U\right) \leq \tau$ ▷ Convergence
Return: $\hat{X}, \hat{Y}, \hat{\theta}$

This modification brings two additional benefits. First, it enables cross-validation. Without this modification, cross-validation could not be carried out because the original EM algorithm does not handle missing data. The only way to validate the model results, as seen in Shumway and Stoffer (2011) and Cheng and Sabes (2006), would be by way of bootstrapping and hypothesis testing—a validation procedure that yields an empirical distribution of each model parameter and determines the importance of the input variables, but that does not provide any information on the model's predictive skills. Second, the gap filling modification enables the learning algorithm to handle missing data in the instrumental record itself—these missing data points can be replaced by their best available estimates during the learning-reconstruction process.

In principle, the gap filling modification can be implemented to models with higher dimensions (e.g. multi-site models, higher order Markov processes, or higher dimensional state space with different interpretations). There are two potential issues when LDS and its extensions are implemented. Firstly, modelers may face with numerical stability issues in higher dimensions with a lot of missing data, as the computation requires a small amount of matrix inversion (Section 2.3.2). Secondly, equifinality may arise when different parameter sets yield the same goodness-of-fit. In this case, care must be taken when choosing an appropriate model that can be interpreted physically.

2.3.4 Stochastic Streamflow Generation

The LDS model formulated in Section 2.3.1 is a stochastic process model. Once the model's parameters are known, it can be used readily as a stochastic streamflow generator. To do so, one first generates an initial state $x_1 \sim \mathcal{N}(\mu_1, V_1)$. Then, sequentially for each time step $t = 1, ..., T$, the noises $w_t \sim \mathcal{N}(0, Q)$ and $v_t \sim \mathcal{N}(0, R)$ are generated; and x_{t+1} and y_t are computed according to Equations 2.2 and 2.3. This yields one stochastic replicate of the streamflow process and catchment state. The procedure is repeated for the desired number of replicates.

Note that the stochastic replicates generated this way are only associated with one realization of u. As with other stochastic models with exogenous inputs (e.g., linear regression and ARMAX), a hierarchical procedure can be used: one first creates stochastic replicates of u, and, then, for each realization of u, generates replicates of y. When u is the PDSI, generating its stochastic replicates using a time series model can be difficult (Alley, 1984). To alleviate this, one may adopt a nonparametric resampling method, such as the stationary bootstrap (Politis & Romano, 1994).

2.3.5 Experiment Setup

As a basis to gauge the performance of our dynamic model, we created a benchmark reconstruction using principal component linear regression, a well-known method in paleohydrology (cf. Hidalgo et al., 2000; Woodhouse et al., 2006). Specifically, we used a procedure very similar to Woodhouse et al. (2006). First, we performed principal component analysis on the 51 MADA grid points falling within 1,200 km from P.1 station, and retained the highest principal components that cumulatively account for at least 95% of the input variance. We then carried out a backward stepwise linear regression using the retained principal components as predictors, and log-transformed annual streamflow as predictand.

So as to have a fair comparison with the linear regression model, the same input and output variables were used for the LDS model, that is, the principal components selected for the benchmark and log-transformed annual streamflow. For this seminal experiment, we started with a one-dimensional system state for two main reasons: this parsimonious model works well without heavy computational load, and it simplifies the physical interpretation. To further facilitate the physical interpretation, the log-transformed streamflow time series was centralized by subtracting the mean, so that the state x is centralized around zero. We adopted the MATLAB code published by Cheng and Sabes (2006)—available at http://keck.ucsf.edu/~sabes/documents/lds-1.0.tgz.gz—and revised it to accommodate the variant described in Section 2.3.3. Our code is implemented in R and released as a package called *ldsr* (Nguyen, 2020) Since EM is a local optimization algorithm, it may converge to a different maximum likelihood for different initial values of the parameter set $\hat{\theta}_0$. Therefore, we implemented a random search for the initial values of A, B, C, D, Q and R—in the range from -1 to 1—and selected the initial values that yielded the highest likelihood. The *ldsr* package runs this search with parallel processors to speed up the search process. We fixed the value of the algorithm convergence threshold τ equal to 10^{-5} (Shumway & Stoffer, 2011, p.

342) and $x_1 \sim \mathcal{N}(0,1)$. All experiments were carried out on a quad-core Intel i7-6700 CPU @ 3.40 GHz with 32 GB RAM running Microsoft Windows 10.

Both the benchmark and the LDS model were cross-validated with a leave-10%-out cross validation scheme. The reconstruction skills were gauged using coefficient of determination (R^2), normalized root mean squared error (nRMSE), coefficient of efficiency (CE), and reduction of error (RE). The last two metrics are commonly used in dendroclimatology (Fritts, 1976; E. R. Cook et al., 1999; Nash & Sutcliffe, 1970). They are different in the way that the residual sum of squares is normalized. Specifically, let $\mathcal{V}$ be the validation set, then

$$RE = 1 - \frac{\sum_{t\in\mathcal{V}}(y_t - \hat{y}_t)^2}{\sum_{t\in\mathcal{V}}(y_t - \overline{y}_c)^2} \tag{2.28}$$

and

$$CE = 1 - \frac{\sum_{t\in\mathcal{V}}(y_t - \hat{y}_t)^2}{\sum_{t\in\mathcal{V}}(y_t - \overline{y}_v)^2} \tag{2.29}$$

where $\overline{y}_c$ is the mean streamflow in the calibration set, and $\overline{y}_v$ is the mean streamflow in the validation set. Thus, while the Nash-Sutcliffe efficiency is a single metric that measures the model performance on the whole training set, RE and CE separates the model performance into two separate measures: fitness, in the case of RE, and predictive skill, in the case of CE.

Finally, we generated 100 stochastic replicates for the annual streamflow and catchment state following the procedure in Section 2.3.4. Each replicate has the same length as the original reconstruction (406 years). Since our purpose here is only to demonstrate that the LDS model can be used directly as a stochastic streamflow generator, we did not consider the case requiring a stochastic model for the PDSI. Also, 100 replicates should be sufficient to capture the bulk of white noise variability for our demonstration purpose (more replicates may be needed for applications that are sensitive to extreme values).

2.4 Results and Discussion

We first report the results obtained with the LDS model on the instrumental period (1921–2005), and compared them against those provided by a conventional principal component linear regression. Then, we illustrate the reconstructed catchment state and streamflow time series for the entire study period (1600–2005), and discuss their relation with El Niño Southern Oscillation, as well as other climate drivers. Finally, we analyze the stochastic replicates from the LDS model.

2.4.1 Model Performance

On average, LDS has higher R^2 than linear regression while the other metrics have similar mean values for both models (Figure 2.5). LDS, being a more extensive model, may have better skill but may also be more prone to overfitting. On the full data set, however, LDS obtained better streamflow estimation (Figure 2.6). The differences lie mainly where linear regression overestimated or underestimated streamflow for several consecutive years; see for example the periods 1921–1930 and 1948–1954 (Figure 2.6a). This improvement could be attributed to the use of a system state variable—and state-transition equation—in the LDS model. Mathematically, the system state x is a filtered and smoothed version of streamflow; we interpret it as a flow regime state. Thus, the flow regime state x is a quantity that characterizes the annual flow volume compared to the long term mean: $x > 0$ indicates a wet regime, and $x < 0$ a dry regime. The state trajectory revealed regime-like behaviour (cf. Turner & Galelli, 2016): the catchment stayed for years (sometimes decades) in one regime, and then shifted to another regime (Figure 2.6b). By matching the timing of the state trajectory in Figure 2.6b and the streamflow trajectory in Figure 2.6a, one observes that linear regression tended to overestimate streamflow when the catchment was in a dry regime (e.g., 1921-1930) and to underestimate it when the catchment was wet (e.g., 1948–1954), while the LDS model matched observation better. This shows that information about the catchment state may be beneficial.

The catchment state contributes to the streamflow prediction in the LDS model by means of Equation 2.3, which states that the system output is the sum of two terms: the state term Cx_t and the input term Du_t—in other words, streamflow is the result of two components related to the catchment state and exogenous inputs. Given this relationship, the modified EM algorithm model derived the best combination of the state coefficient C and the input-output coefficients D. As C was found positive (0.085), a quantity of $|Cx_t|$ was added to (subtracted from) the input term Du_t when $x_t > 0$ ($x_t < 0$). But this increase (decrease) did not lead to overestimation (underestimation) because the algorithm derived the input coefficients D that have the same signs, but smaller magnitude, than the linear regression coefficients β (Table 2.2). Consequently, the LDS model was able to account for the situations in which the catchment is still wet (dry) following a previous wet (dry) year, although the PDSI for that particular year may not be high (low).

Table 2.2. Comparing exogenous input coefficients for linear regression and linear dynamical systems (LDS) models.

Principal components	β^a	D^b
PC1	-0.0317	-0.0282
PC3	-0.0271	-0.0212
PC6	-0.0615	-0.0513
PC9	-0.1265	-0.1142
PC12	-0.1122	-0.0830

[a] Linear regression coefficients (Equation 2.1)
[b] LDS coefficients (Equation (2.3)

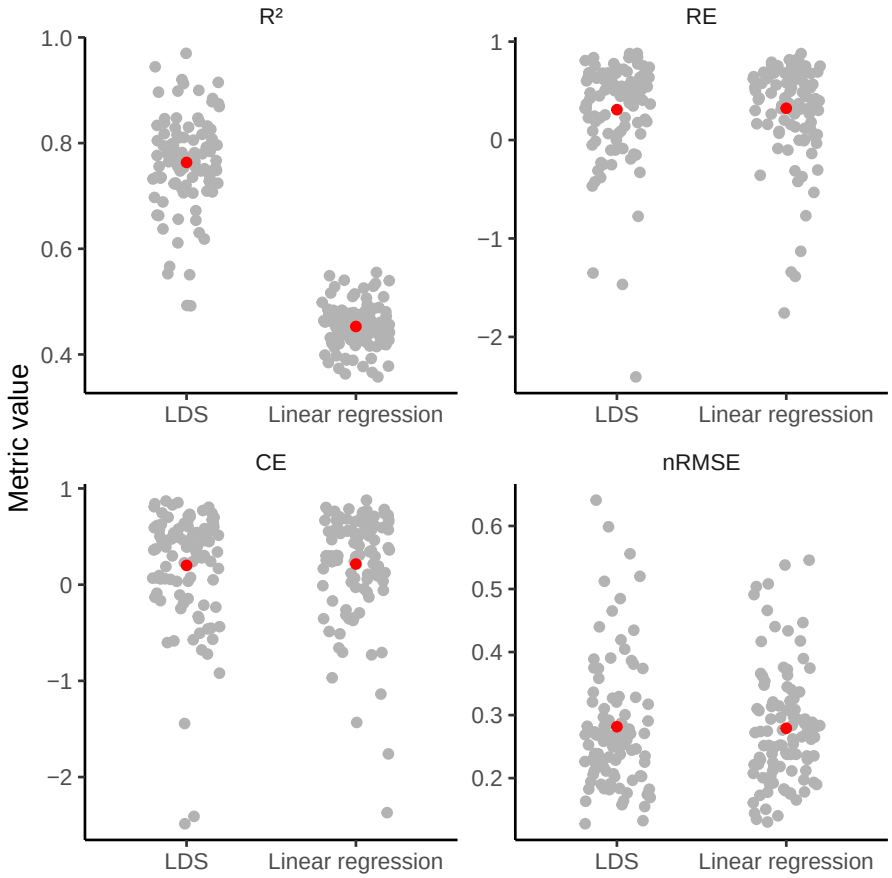

Figure 2.5. Distribution of performance scores in cross-validation runs obtained by linear regression and linear dynamical systems (LDS) models. Gray dots represent the value of each score obtained during the validation runs; red dots represent the average value of the scores across all runs. R^2, RE, CE, and nRMSE denote the coefficient of determination, reduction of error, coefficient of efficiency, and normalized root mean squared error, respectively.

Residual analysis (Figure 2.7) showed that the assumption of independent Gaussian noise was not violated in either models. However, large deviations from Gaussian were observed in both positive and negative tails for the linear regression residuals. For positive residuals (overestimation), the two points of large deviation corresponded to the years 1931 and 1992, during both of which the catchment was in a very dry flow regime (Figure 2.6b). For negative residuals (underestimation), the two points of large deviation corresponded to the years 1973 and 2005, during both of which the catchment was in a very wet flow regime (Figure 2.6b). These large deviations were much less apparent in the LDS results where the flow regime was taken into account, although one may observe that when residuals are transformed from the log space back to the original streamflow space, the deviation is still large for year 2005. Thus, residual analysis further corroborates that catchment dynamics should not be neglected in streamflow reconstruction.

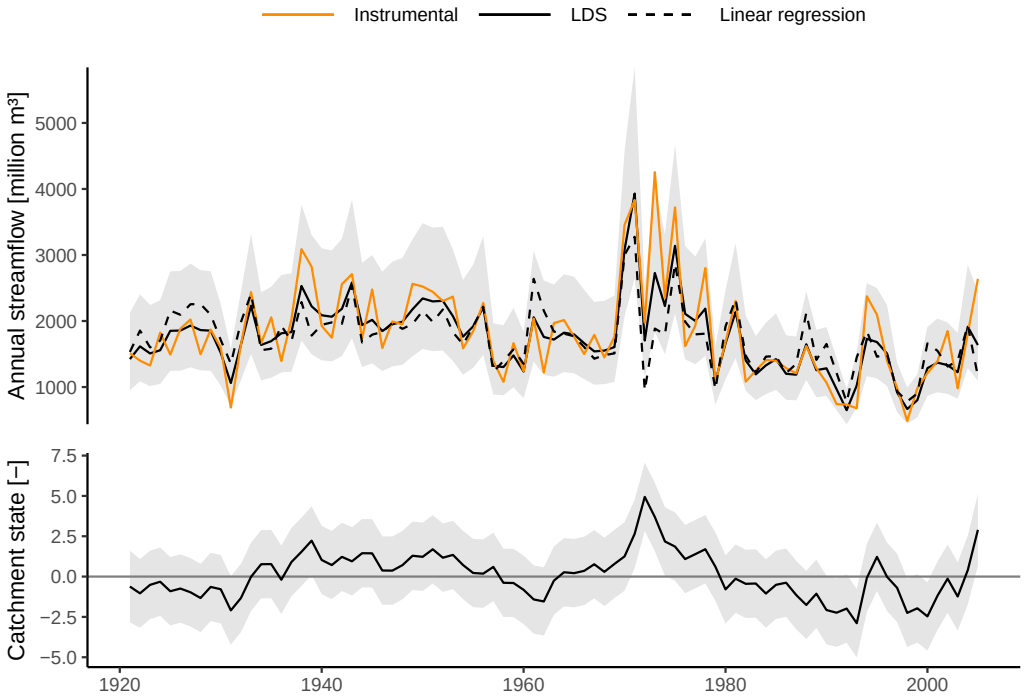

Figure 2.6. Results of the linear dynamical systems (LDS) model in the instrumental period: (a) Reconstructed streamflow, plotted with 95% confidence interval, compared with the instrumental time series and the results from a benchmark linear regression model (Section 2.3.5); b) Trajectory of the system state (flow regime) with 95% confidence interval. LDS generally provided higher streamflow estimates during periods of high flow regime (1935–1955, 1968–1978), and lower streamflow estimates during periods of low flow regime (1921–1935, 1980–1995).

2.4.2 A Reconstructed Hydrological History of the Ping River

Results revealed a history of droughts, floods and regime shifts in the Ping River over the last four centuries (1600–2005). The LDS model and linear regression yielded similar results in normal years (i.e., when the flow regime is about zero), but the LDS model provided lower streamflow estimates in dry years and higher streamflow estimates in wet years (Figure 2.8a). Most importantly, the LDS model provided a drier picture than what was seen in linear regression results, especially during the low flow periods. This result may have important implications to water management—for instance, in the form of more conservative operating policies for the Bhumibol Reservoir.

The reconstructed flow regime shows different patterns of regime shift over time (Figure 2.8b). At first, the flow regime shifted infrequently in the 17th century; there were four main wet and dry epochs that lasted more than a decade (an epoch is a period where streamflow persists in the same regime). The flow regime then shifted more rapidly in the 18th and 19th century, where each wet or dry epoch lasted only a few

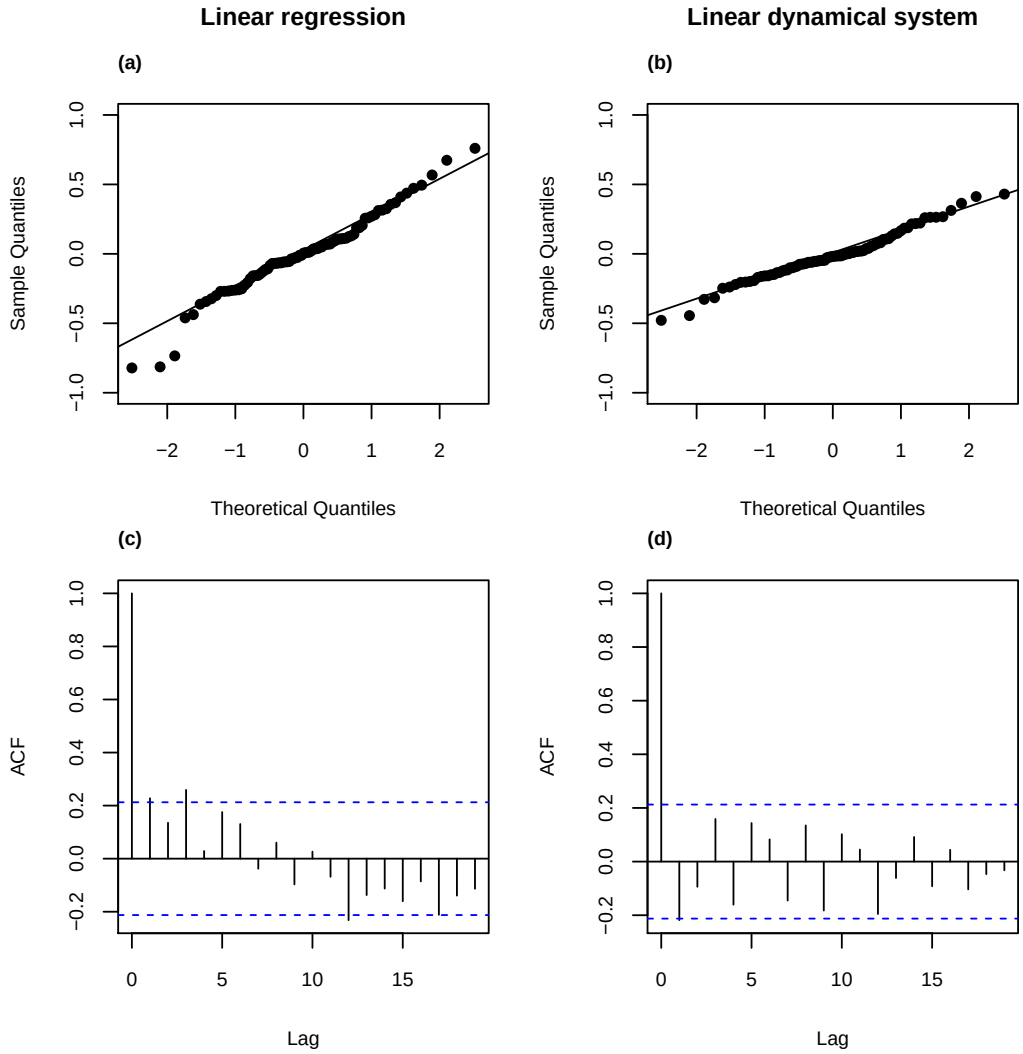

Figure 2.7. Residual analysis results for the linear regression (left column) and linear dynamical systems (right column) models. The analysis was based on the log-transformed streamflow. Panels (a) and (b) show the quantile-quantile plots of the residuals compared to Gaussian distributions; both models' residuals are close to Gaussian, but larger deviations are observed in the tails for linear regression. Panels (c) and (d) show the autocorrelation function (ACF) of the residuals; no significant autocorrelations are observed.

years. The pattern of regime shift is most varied in the 20th century: there were prolonged wet and dry epochs of decadal to bi-decadal scales (similarly to the 17th century). However, the flow regime fluctuated more vigorously than the previous three centuries. As a result, the last century contains both the wettest period (including the wettest year) and the driest year on record. During the wettest period (1966–1979), two consecutive strong La Niña events occurred, and the driest year (1998) corresponded to a very strong El Niño event. We discuss this correspondence further in Section 2.4.3.

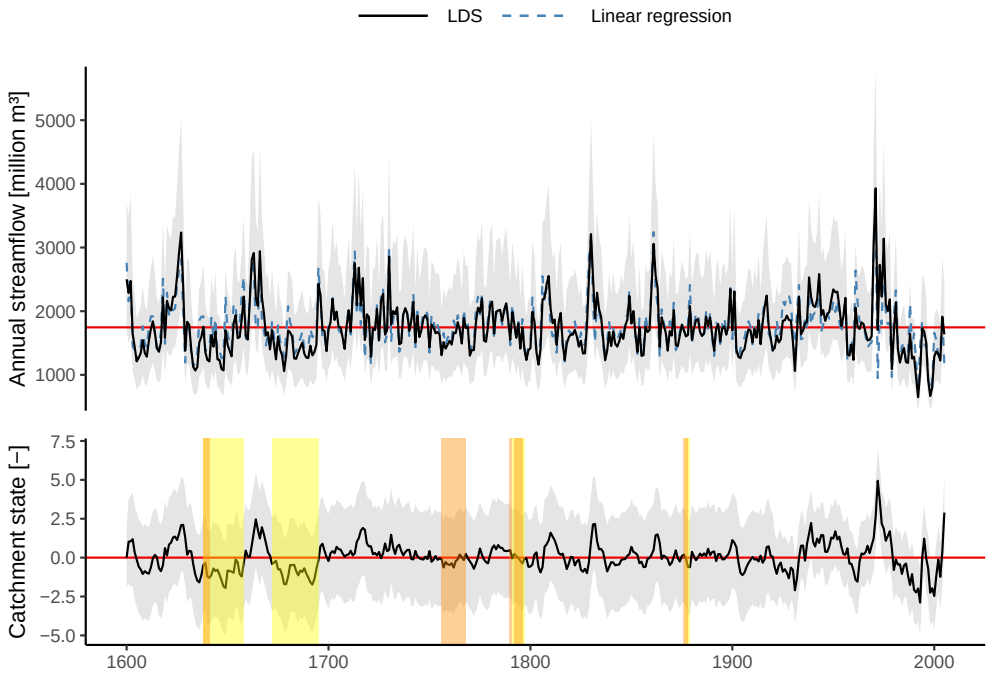

Figure 2.8. Full reconstruction results: a) Reconstructed streamflow, compared with linear regression; b) Flow regime state trajectory. The grey bands are the 95% confidence intervals. The red lines indicate the mean values of the full reconstruction. In panel b, the orange bands are the megadroughts discussed in E. R. Cook et al. (2010a), namely the Ming Dynasty Drought (1638–1641), the Strange Parallels Drought (1756–1768), the East India Drought (1790, 1792–1796) and the Victorian Great Drought (1876–1878). The yellow bands are the dry epochs revealed by the flow regime state variable in the paleo period (a dry epoch is a period of consecutive negative flow regime).

The LDS model results are in agreement with the MADA, in that all four Asian megadroughts in the last millennium each impacted northern Thailand (E. R. Cook et al., 2010a). These are: the Ming Dynasty Drought (1638–1641), the Strange Parallels Drought (1756–1768), the East India Drought (1790, 1792–1796) and the Victorian Great Drought (1876–1878). But, more interestingly, while the MADA provided a geographical footprint of these droughts, our reconstruction provided more insights pertinent to the Ping River (Figure 2.8). The Ming Dynasty Drought seems to have triggered, or at least contributed, to a prolonged dry epoch in the Ping. By 1638, the Ping River was coming out of a short dry epoch. The occurrence of the Ming Dynasty Drought then coincided with three years of declining streamflow, which set the Ping back to a dry epoch that took two decades to vanish. Throughout this drought, streamflow stayed below the mean level. The Strange Parallels Drought was in the middle of several decades where streamflow was mostly at or below normal, and the flow regime was mostly around zero. Anchukaitis et al. (2011) suggested that the Strange Parallels is closely related to the Indian Ocean Dipole (IOD), and we found that the same relationship holds for the Ping River. Comparing the state trajectory in Figure 2.8b with page 40 in Anchukaitis et

al. (2011), we observe that the first half of the drought, where flow regime was around zero, is consistent with a negative phase in the IOD, and the spike at the end is consistent with a brief positive phase in the IOD. The flow regime history suggested that the Strange Paralells was hydrologically mild, yet it coincided with a tumultuous part of Southeast Asia's history (Lieberman, 2003; E. R. Cook et al., 2010a), indicating that the socio-economic damage of this drought may have been more serious than its hydrologic impact. The East India Drought coincided with a dry epoch in the Ping River, but this drought seemed to have a lesser impact in Thailand than other megadroughts. The Victorian Great Drought was similar to the Ming Dynasty Drought, in that it also set into motion a dry epoch. The MADA indicated a meteorological drought in the whole Southeast Asia between 1876–1878 while the flow regime indicated a hydrological drought in Thailand between 1878–1886. This suggests that the catchment may have seen a transition from a meteorological drought to a hydrological one. The flow regime also indicated a major drought between 1687–1696 that was not identified as a megadrought in E. R. Cook et al. (2010a), suggesting that this drought was more localized to Thailand. It should be noted that the PDSI is a meteorological drought index (W. C. Palmer, 1965; Alley, 1984) that does not always reflect hydrological droughts (Mishra & Singh, 2010). Hence, droughts identified by the PDSI and those identified by the flow regime may have similarities and differences. This implies that a regional drought footprint and a local streamflow reconstruction can complement each other to provide better understanding, as we demonstrated here.

The LDS results also identified multiple wet epochs in the three centuries preceding the instrumental records, with several pluvial years having flow comparable to the highest ones in the instrumental period (Figure 2.8). Notably, a prolonged wet epoch occurred between 1659–1672, consistent with a period of seven floods circa 1658 $\pm$ 7 years, identified in a paleoflood study using analyses of river sediments [*Wasson, 2017, personal communication*]. The same study also identified a major flood in 1831, consistent with the wet epoch between 1830–1838 as shown in the state trajectory. This result somewhat reflects the flood generation mechanism of the Ping River, where floods are due to heavy rainfall events occurring over a wet catchment. It should also be noted that the sediment study identified peak discharge events, while we reconstructed annual streamflow. Maximum annual flow volume and peak discharge may not necessarily occur in the same year, but our results showed that the catchment stayed in the wet regime several years after a major flood.

2.4.3 Modes of Streamflow Variability

To characterize the most important temporal modes of variability contained in the reconstructed streamflow time series, we carried out a wavelet analysis using the Morlet wavelet (Roesch & Schmidbauer, 2014). We also applied the same technique to the reconstructed Sea Surface Temperature (SST) anomalies in the Eastern Pacific (Tierney et al., 2015). As shown in Figure 2.9, reconstructed streamflow shows a mode of variability that coincides with the frequency of the El Niño Southern Oscillation (ENSO)—about 2 to 7 year—during the 17th century, early 18th century and, intermittently, 20th century. This result is consistent with previous studies for Thailand, Vietnam, and Southeast Asia (i.e., Buckley et al. (2007b), Sano et al. (2009), and Räsänen et al. (2016)),

which suggest that positive ENSO anomalies result in reduced PDSI, and, hence, reduced precipitation. Yet, our results do not indicate a perfect match between interannual variability in SST anomalies and reconstructed streamflow. This may be explained by the fact that SST anomalies in the the Eastern Pacific do not always lead to ENSO events (Dunbar et al., 1994; Buckley et al., 2007b). Furthermore, Singhrattna et al. (2005) reported that the effect of ENSO on the Thailand summer monsoon exhibits time dependence. In particular, the same authors showed that the relationship between ENSO and Thailand rainfall became stronger after the 1980s; this might explain the steady ENSO-like temporal mode of variability we observe for the reconstructed streamflow during that period.

Frequency analysis results for the reconstructed streamflow also show features of interdecadal variance in the 17th, 19th, and 20th century, which are consistent with the

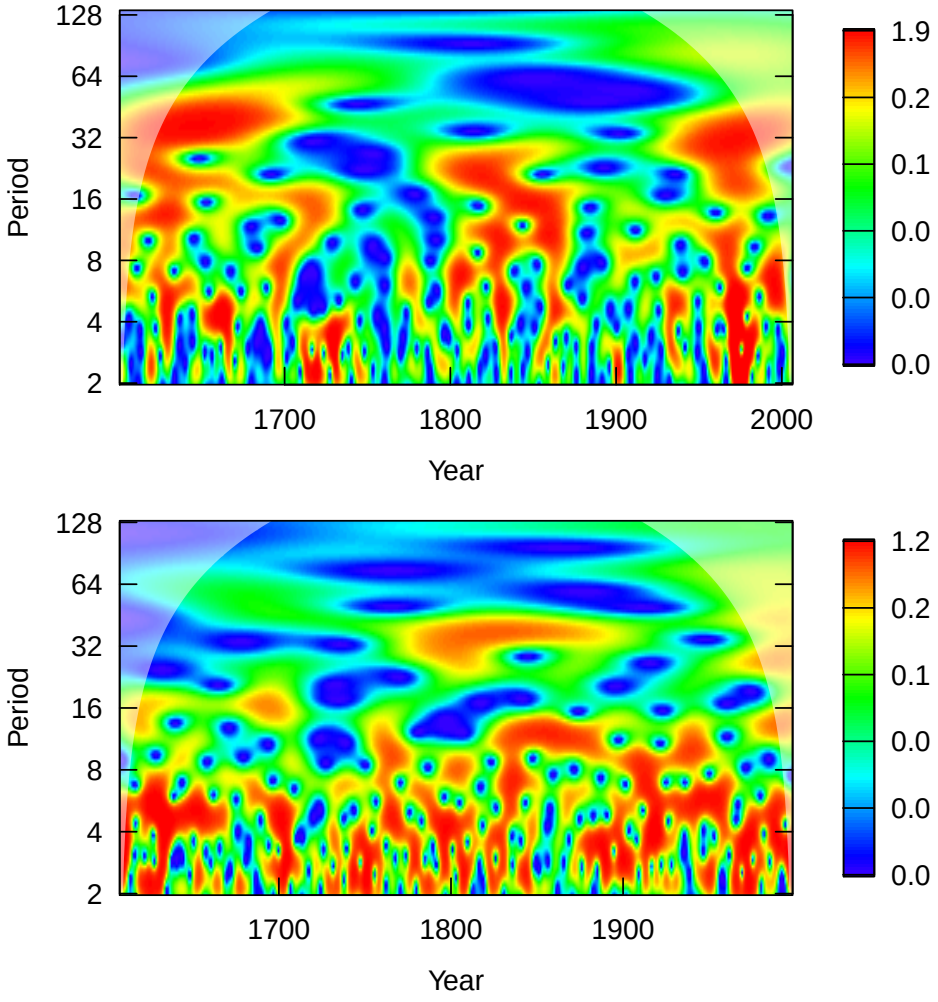

Figure 2.9. Wavelet analysis of a) Reconstructed streamflow; and b) Reconstructed Eastern Pacific SST anomalies (Tierney et al. (2015)). The colour bars indicate the wavelet power. (Values in the fainted region are outside of the cone of influence and should not be interpreted.)

prolonged wet and dry epochs described in Section 2.4.2. As noted in previous studies (Buckley et al., 2007b; Sano et al., 2009), these results indicate that other climate drivers may cause decadal streamflow variability in region. For instance, Sano et al. (2009) found a significant positive correlation between tree-ring reconstructions in Vietnam and SST in the northern Pacific Ocean, suggesting a possible link with the Pacific Decadal Oscillation (Mantua & Hare, 2002). The Indian Ocean Dipole (Saji et al., 1999) have also been found to be related to floods and droughts in Southeast Asia (Delgado et al., 2012; Räsänen & Kummu, 2013). This phenomenon may also influence streamflow variability in the Ping River, as indicated in our discussion on the Strange Parallels Drought (Section 2.4.2).

2.4.4 Stochastic Replicates

Figure 2.10 shows the stochastic ensemble generated by simulating the LDS model with different realizations of the noise processes w and v according to Equations 2.2 and 2.3. The ensemble contains a wide range of high and low values enveloping the reconstructed time series of both streamflow and flow regime. Interestingly, the envelopes follow the shape of the reconstructions closely: large floods and droughts tend to occur at the same time in the replicates as in the reconstruction. Keeping in mind that these replicates were generated with the historical exogenous input, i.e., the same inputs that drove the variability of the reconstructions (Section 2.3.4), the ensemble envelopes indicate a strong input-driven dynamics: while the noise processes (which represent factors unaccounted for in the model) may lead to different streamflow and flow regime values, they cannot mask away the general trajectory dictated by the exogenous inputs. This implies that the shape of the reconstructed trajectories and the corresponding wet and dry states are robust.

2.5 Conclusions

In this chapter, we contributed a technique for streamflow reconstruction based on the state-space representation of a discrete, linear dynamical system, which was learned using a novel variant of the Expectation-Maximization algorithm. The use of a state-space representation yields two key advantages: it estimates the trajectory of the catchment state during the paleo and instrumental periods, and it accounts for the effect of both catchment state and climate proxies on the streamflow generation process. The technique was tested to reconstruct 406 years of annual streamflow for the Ping River, northern Thailand, using the Monsoon Asia Drought Atlas gridded PDSI data set (E. R. Cook et al., 2010a) as the paleoclimate proxy. Our reconstruction identified several prolonged pluvials and droughts in the paleo period. Somewhat differently from most previous works, we found that the instrumental record contains both the wettest period 1965–1979) and the driest year (1998). Our results are aligned with earlier works (e.g., Ho et al., 2015; Tozer et al., 2016) in that flood and drought analyses based on reconstructed streamflow may yield a different picture from similar analyses using instrumental data. Therefore, as seen with other regions (E. R. Cook et al., 2010b; Kiem et al., 2016; Johnson et al., 2016), there is a need of more reconstruction studies in Southeast Asia to better understand these natural hazards.

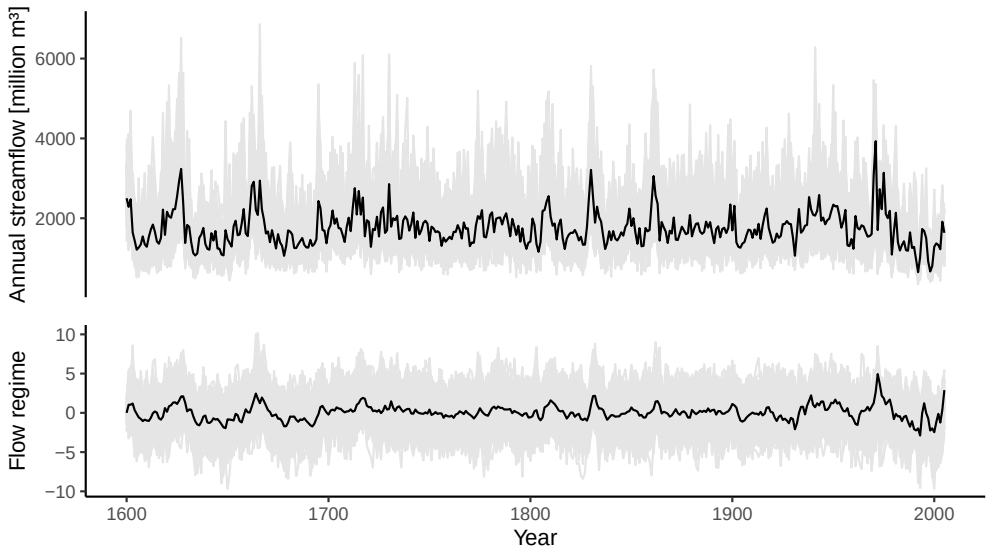

Figure 2.10. Stochastic replicates generated from the LDS model. The black lines are the original reconstructions (Section 2.4.2), and the gray lines are the 100 stochastic replicates, which were generated by adding noise to the original reconstruction according to Equations 2.2 and 2.3.

The model's reconstruction in the instrumental period is reliable, supporting the findings by Ho et al. (2016) that a paleodrought record can be used to reconstruct streamflow, and by Watson and Luckman (2005) that gridded PDSI data sets are a rich source of information to investigate hydroclimatic variability. A qualitative comparison between LDS and linear regression on pluvials and droughts suggests that it is important to account for catchment dynamics, especially in systems characterized by complex streamflow generation processes. Additionally, our linear dynamical system model has several desirable features. (i) The reconstructed trajectory of the state variable provides more insights about the catchment's history than the reconstructed streamflow alone. For instance, we have shown that the model's state variable reveals regime-like behavior of streamflow. (ii) The Expectation-Maximization algorithm used to learn the model is computationally efficient, and does not require any assumption about serial dependence. (iii) The model can be readily used as a stochastic streamflow generator, and it is extendable to multi-site applications.

A natural extension of our technique is the identification of a nonlinear dynamical system model, in which the state and output equations are nonlinear. In this case, as suggested by Roweis and Ghahramani (2001), the Kalman smoother in the E-step needs to be replaced by an extended Kalman smoother, and the global optimizer in the M-step can no longer be determined analytically. Such model is thus more computationally expensive, but it may yield better results—particularly in catchments that present strong nonlinearities associated to the streamflow generation process. The benefit and cost of such nonlinear models should be investigated. Another possible extension is to use a

multi-dimensional state vector, as only one state variable was used here. It is perceivable that multiple state variables may contain more information or improve model performance; how to interpret them remains an open question. A third extension could be to explore models in which the parameters A, B, C, D, Q and R change over time. Such time-varying models may be used to account for changes in catchment characteristics, e.g., due to changes in land use/ land cover. When applying LDS and its extensions, modellers should be aware of potential issues such as numerical stability and model equifinality.

Perhaps the most relevant application of this work that should be the topic of immediate research is to transfer the added understanding of catchment dynamics to water management practices, such as reservoir operation models. Recently, Turner and Galelli (2016) and Ng et al. (2017) have shown that regime-like behaviour in streamflow time series contributes to the sub-optimality of reservoir operating policies derived with conventional optimization methods; on the flip side, better operating policies can be obtained by incorporating a regime state variable into reservoir operations. In addition, robust operating policies require longer streamflow records, since more training data are likely to provide more robust operating policies. Reconstruction studies that model regime state, such as this work, address both needs.

The encouraging results and the desirable features of the LDS model suggest that it can replace linear regression in future streamflow reconstruction studies. Most importantly, the model's regime state, not available in conventional methods, may add value to downstream applications such as reservoir operations studies. Through the findings in this work, not only has the values of streamflow reconstruction been strengthened, but its potential applications have also been widened.

Chapter 3

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

Preprint

Nguyen, H. T. T., Galelli, S., Xu, C., & Buckley, B. M. (2020a). Multi-Proxy, Multi-Season Streamflow Reconstruction with Mass Balance Adjustment. *Earth and Space Science Archive*. https://doi.org/10.1002/essoar.10504791.1 (*submitted to Water Resources Research*).

3.1 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 that affected irrigation. 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. That is because reconstructions often target specific components of the hydrograph that best correlated with tree ring proxies. Perhaps most commonly, reconstructions

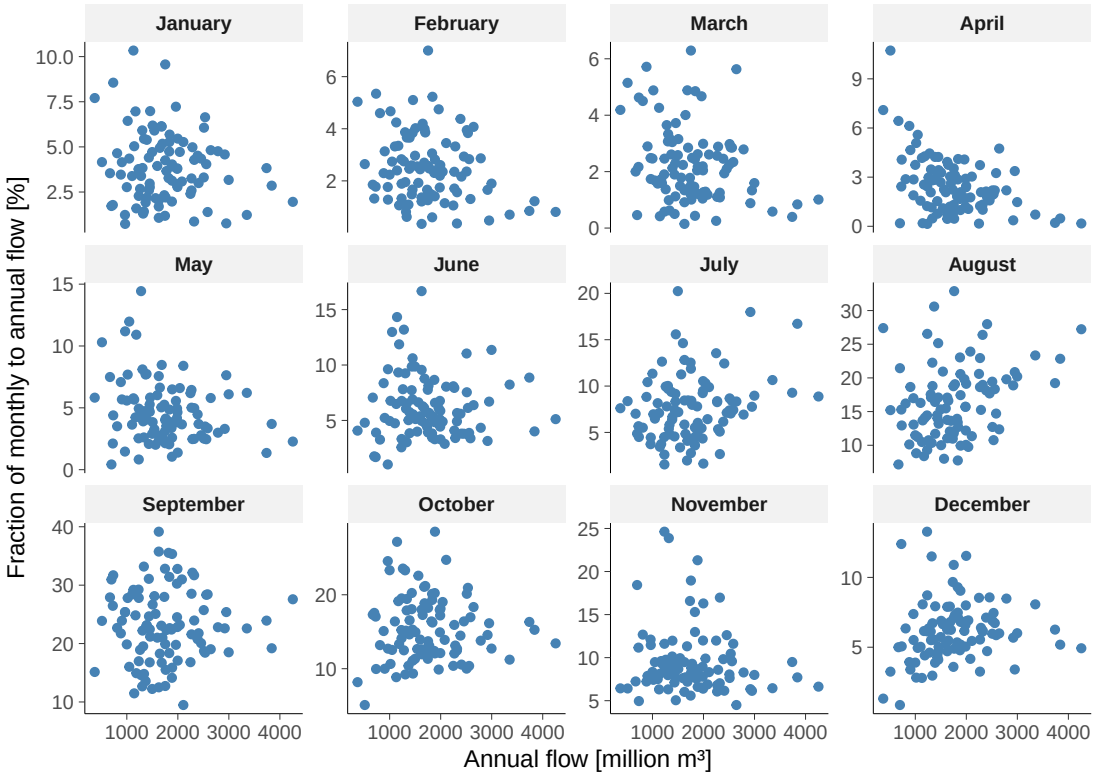

Figure 3.1. Ratio between monthly flow and annual flow for each month at station P.1, plot against the annual flow volume. In each month, for a similar range of annual flow, there can be a wide range of monthly flow ratios. This demonstrates that statistical disaggregation based on the constant ratio assumption may not always hold.

from ring width target the growth season (e.g., D'Arrigo et al., 2011a; Güner et al., 2017). Another example is given by reconstructions using tree ring cellulose $\delta^{18}O$ and targeting peak flow (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., Rao et al., 2018; Nguyen & Galelli, 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 towards this goal have relied on statistical disaggregation, assuming some statistical relationships between the sub-annual and annual flows (Prairie et al., 2008; Sauchyn & Ilich, 2017). These assumptions are reasonable but not always valid (Figure 3.1). 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. Therefore, instead of disaggregation, we propose to use multiple proxies to simultaneously reconstruct sub-annual (e.g., seasonal) and annual flows. Two challenges arise. How to combine proxies optimally for different targets? And how to ensure that the seasonal flows add up to the annual flow, i.e., 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 3.3.1), and proxy combination is optimized with an automatic input selection scheme (Section 3.3.2). 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 $\delta^{18}O$ chronologies from Southeast Asia (Section 3.2). This work is one step closer towards operational usability of streamflow reconstruction in water resources management.

3.2 Study Site and Data

3.2.1 The Southeast Asian Dendrochronology Network

Over the past three decades, an extensive network of tree ring chronologies have been developed in mainland Southeast Asia. This network has been instrumental in improving our understanding of Southeast Asia's hydroclimate and history. For example, tree ring data from Thailand and northern Vietnam (Buckley et al., 2007b; Sano et al., 2009) revealed a multidecadal drought, what is later known as the Strange Parallel Droughts (E. R. Cook et al., 2010a), 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 14^{th} and 15^{th} centuries to the demise of the Angkor Civilization (Buckley et al., 2010; Buckley et al., 2014). These findings are among many insights that the Southeast Asian Dendrochronology Network has brought forth.

In this work, we use twenty tree ring chronologies from Vietnam, Laos, Cambodia, Thailand, and Myanmar (Figure 3.2). The chronologies at Kirirom, Petchaburi, Pha Taem, and Wiang Haeng are published here for the first time. The details of the chronologies are provided in Table 3.1. The common period of most chronologies in our network is 1748–2005 (Figure B.1), and is the same as the time span of our $\delta^{18}O$ 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 B.2). 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 3.3.1, 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.

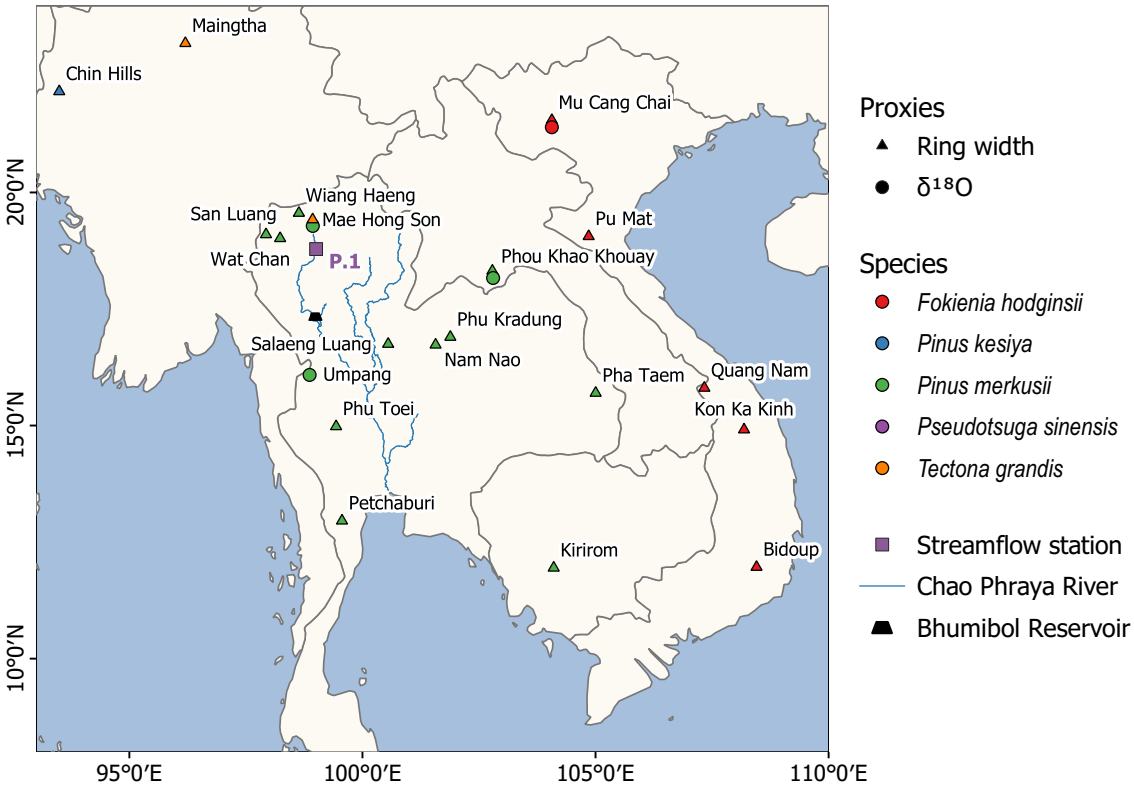

Figure 3.2. Map of the study region, showing the Chao Phraya River, the proxy network, and the target streamflow station.

3.2.2 Tree ring cellulose stable oxygen isotope

We use four tree ring cellulose oxygen isotope ($\delta^{18}O$) chronologies that were developed in Laos, Thailand, and Vietnam over the past decade (Figure 3.2 and Table 3.2). $\delta^{18}O$ exhibits strong mechanical and statistical relationship with the hydroclimate (C. Xu et al., 2011; Sano et al., 2012), and has been used to reconstruct wet season precipitation in the region (C. Xu et al., 2015; C. Xu et al., 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 works support recent literature (Treydte et al., 2006; G. Xu et al., 2019) that $\delta^{18}O$ has indeed moved beyond "potential" (Gagen et al., 2011), and earned its place as a practical, valuable paleoclimate proxy.

3.2.3 Streamflow

The Chao Phraya River Basin covers 30% of Thailand's area, and is home to about 25 million people. Thailand's largest reservoir, the Bhumibol (active capacity 9.7 km^3), lies on the Ping River tributary. Reliable operations of this reservoir require accurate assessment of inflow availability, on both inter- and intra-annual scales.

Table 3.1. 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 (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 et al. (2007b)
Maingtha	96.20	23.20	*Tectona grandis*	D'Arrigo et al. (2011b)
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 et al. (2007a)
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. (2010a)
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. (2010a)
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.

Streamflow station P.1 is located upstream of Bhumibol Reservoir. P.1 has the longest and most complete record in Thailand: daily data are available from April 1921–present. 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, thereby interfering with the proxy-streamflow relationship. Therefore, we naturalized the streamflow data from 1985. The naturalization process is described in Appendix B.2.3. 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) (Appendix B.2.2). To match the proxies' time span, we finally used the

Table 3.2. Metadata of $\delta^{18}O$ 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)

streamflow data from November 1921 to October 2005.

3.2.4 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. Negative lags account for the case when trees use stored carbon from previous years, and positive lags for the case when the catchment's runoff processes are slower than precipitation inputs (Stockton & Jacoby, 1976; Meko et al., 2007). For robustness, we repeated the correlation analysis 1,000 times using the stationary bootstrap (Politis & Romano, 1994). In the following discussion we refer to the median bootstrap correlations (Figure 3.3).

Among the ring width sites, there are multiple correlation patterns (Figure 3.3a): some sites such as Chin Hills and Phu Toei correlate positively, while others (e.g., Phou Khao Khouay) correlate negatively. Peculiarly, the Mae Hong Son site displays significant negative correlation at $l = -2$ but significant positive correlations at $l = 0$ and $l = 2$. Five sites do not correlate with streamflow at all. These various patterns suggest that the ring width–streamflow relationship is complex and "noisy". A large number of sites are thus required to extract the strongest signals.

Unlike ring width, $\delta^{18}O$ displays more consistent correlation patterns (Figure 3.3b): 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 $\delta^{18}O$ chronologies may contain stronger climate signals than do ring width chronologies (C. Xu et al., 2019; Gagen et al., 2011).

In general, we observe that ring width tends to correlate more strongly with dry season flow than with wet season flow. Conversely, $\delta^{18}O$ 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 (Appendix B.3). 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.

3.3 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 $\delta^{18}O$ at lag 0 (Figure 3.3). For 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 3.3.1), 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

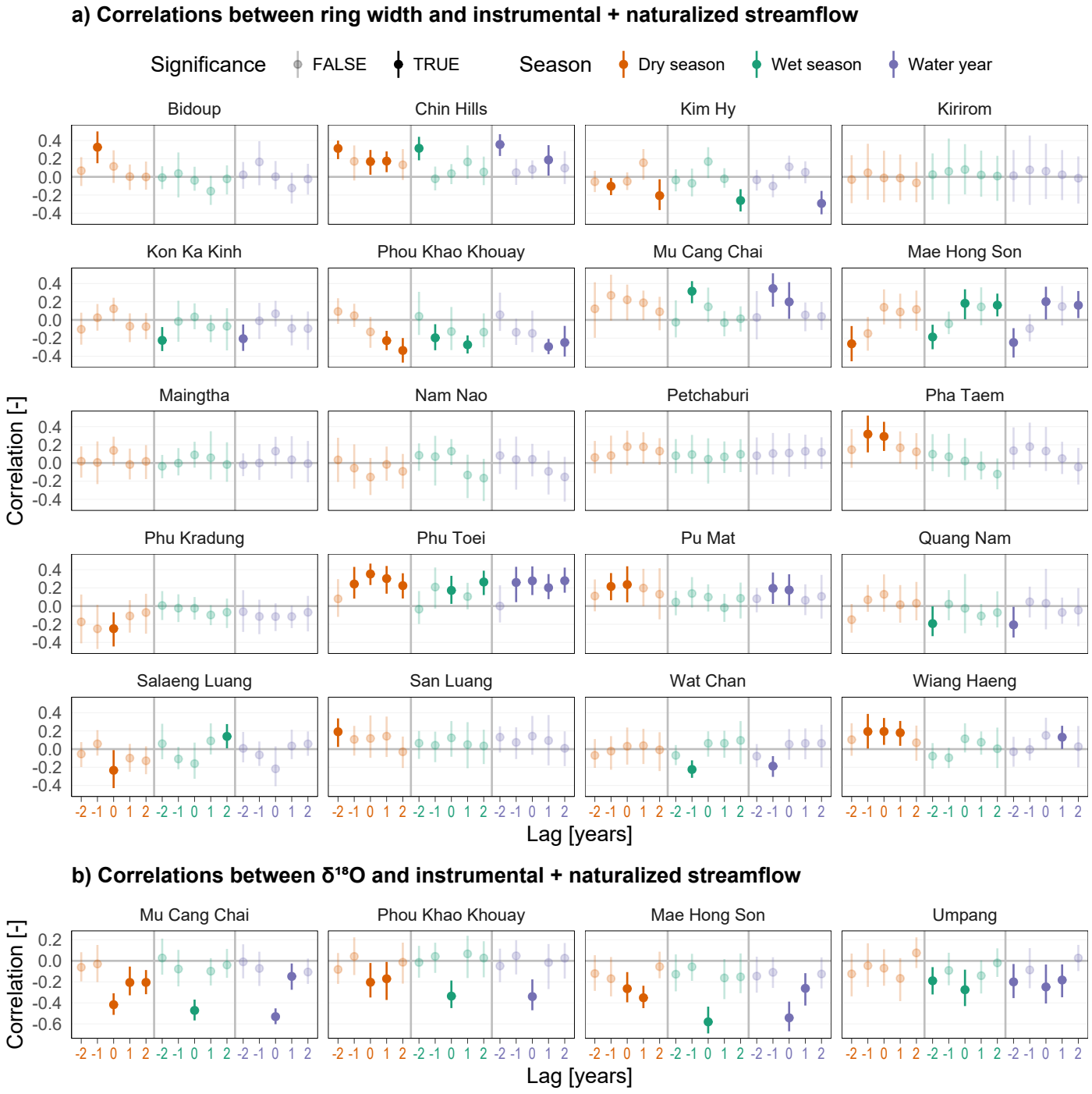

Figure 3.3. Streamflow–proxy correlations. The error bars show the 5th–95th bootstrapped empirical quantiles obtained from 1,000 replicates, using the stationary bootstrap (Politis & Romano, 1994). The dots indicate the medians. Lag l denotes correlations between proxy at year t and streamflow at year $t + l$.

(Section 3.3.2), 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).

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 (Hidalgo et al., 2000; Coulthard et al., 2016). 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 $g(.)$:

$$\mathbf{X_d} = g(\mathbf{U_d}) \tag{3.1}$$

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

$$\mathbf{X_w} = g(\mathbf{U_w}) \tag{3.2}$$
$$\mathbf{X_q} = g(\mathbf{U_q}) \tag{3.3}$$

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:

$$\mathbf{y_d} = \mathbf{X_d}\boldsymbol{\beta}_d + \boldsymbol{\varepsilon}_d \tag{3.4}$$
$$\mathbf{y_w} = \mathbf{X_w}\boldsymbol{\beta}_w + \boldsymbol{\varepsilon}_w \tag{3.5}$$
$$\mathbf{y_q} = \mathbf{X_q}\boldsymbol{\beta}_q + \boldsymbol{\varepsilon}_q \tag{3.6}$$

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

$$\mathbf{y} = \begin{bmatrix} \mathbf{y_d} \\ \mathbf{y_w} \\ \mathbf{y_q} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{X_d} & & \\ & \mathbf{X_w} & \\ & & \mathbf{X_q} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_d \\ \boldsymbol{\beta}_w \\ \boldsymbol{\beta}_q \end{bmatrix}, \quad \text{and} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_d \\ \boldsymbol{\varepsilon}_w \\ \boldsymbol{\varepsilon}_q \end{bmatrix}. \tag{3.7}$$

Equations 3.4–3.6 can then be converted to a more compact form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{3.8}$$

Equation 3.8 has the canonical form of linear regression. It can be solved as a least-squares problem:

$$\min_{\boldsymbol{\beta}} \quad J_1 = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \tag{3.9}$$

yielding the solution

$$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{3.10}$$

Solving Equation 3.8 is equivalent to solving Equations 3.4–3.6 simultaneously. The three regression problems in Equations 3.4–3.6 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,

$$\mathbf{X_d}\boldsymbol{\beta}_d + \mathbf{X_w}\boldsymbol{\beta}_w = \mathbf{X_q}\boldsymbol{\beta}_q. \tag{3.11}$$

But, Equation 3.11 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 3.9 a penalty term that is based on the differences (δ) between the LHS and the RHS of Equation 3.11.

$$\delta = \mathbf{X_d}\boldsymbol{\beta}_d + \mathbf{X_w}\boldsymbol{\beta}_w - \mathbf{X_q}\boldsymbol{\beta}_q. \tag{3.12}$$

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

$$\delta_t = \log\Big(\exp(\mathbf{x_{d,}}_t\boldsymbol{\beta}_d) + \exp(\mathbf{x_{w,}}_t\boldsymbol{\beta}_w)\Big) - \mathbf{x_{q,}}_t\boldsymbol{\beta}_q \qquad \forall t = 1, ..., T. \tag{3.13}$$

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

$$\min_{\boldsymbol{\beta}} \quad J_2 = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \lambda\boldsymbol{\delta}'\boldsymbol{\delta} \tag{3.14}$$

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.

Lemma 3. *Without flow transformation, the analytical solution of Equation 3.14 is*

$$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X} + \lambda\mathbf{A}'\mathbf{A})^{-1}\mathbf{X}'\mathbf{y} \tag{3.15}$$

where $\mathbf{A} = \begin{bmatrix}\mathbf{X_d} & \mathbf{X_w} & -\mathbf{X_q}\end{bmatrix}$.

The proof is provided in Appendix B.4.

When log-transformations are involved, δ is not linear, and Equation 3.14 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()`.

3.3.2 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 3.14). If the number of inputs is small, we can exhaustively search all possible subsets and choose the one that yields the minimum penalized least square 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 i^{th} input is selected. For any given $\mathbf{p}$, i.e, 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:

$$d_i = \begin{cases} 1 & \text{if proxy } i \text{ is used for the dry season} \\ 0 & \text{otherwise} \end{cases} \qquad i = 1, ..., n_d. \tag{3.16}$$

So, where $d_i = 1$, we take the i^{th} 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 3.3.3), 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

$$\min_{\mathbf{p}} \; f(\mathbf{p}). \tag{3.17}$$

We solve Equation 3.17 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 Appendix B.5.

3.3.3 Model Assessment

We set up a reconstruction experiment involving two models: Model 0 runs without the mass balance adjustment ($\lambda = 0$ in Equation 3.14) and Model 1 has the adjustment ($\lambda = 1$). Other than the different values for λ, both models are trained exactly the same way, following Sections 3.3.1 and 3.3.2.

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) (Nash & Sutcliffe, 1970; Fritts, 1976). All metrics are calculated over 50 cross-validation runs.

Following Nguyen et al. (2020b), we adopt a leave-25%-out cross-validation scheme, where each hold-out chunk is contiguous. 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. The 50 repetitions provide a distribution for each skill metric, allowing more robust estimation of the mean skill score. More importantly, the distributions enable us to assess the statistical significance of skills. For example, a reconstruction is considered statistically skillful with respect to CE at $\alpha = 0.1$ if the probability of negative CE is less than 0.1.

3.4 Results

3.4.1 Reconstructions

For Model 0's dry season and annual reconstructions, all metrics are at least 0.40, and the reconstructions match their targets closely (Figure 3.4). Furthermore, these reconstructions are statistically skillful at $\alpha = 0.1$. Conversely, Model 0's wet season reconstruction is not statistically skillful. Although the mean RE and CE are positive (RE = 0.35, CE = 0.23), these scores vary widely over the cross-validation runs (Figure 3.5), suggesting that the wet season reconstruction is sensitive to training data. The large variability of skills is also consistent with the high variability of streamflow (Figure B.3). These observations suggest that there may be nonlinearity in the streamflow–proxy relationships at the extremes. In future studies, nonlinear reconstruction models (e.g., Nguyen & Galelli, 2018) could be incorporated to address this problem.

Model 1, with the penalty term, makes visible adjustments to the seasonal reconstructions but minimal changes to the annual one for the instrumental period (Figure 3.4). Dry season skills slightly decrease, wet season's RE and CE increase, and annual skills remain almost the same. While the mean skill scores of both models are similar, Model

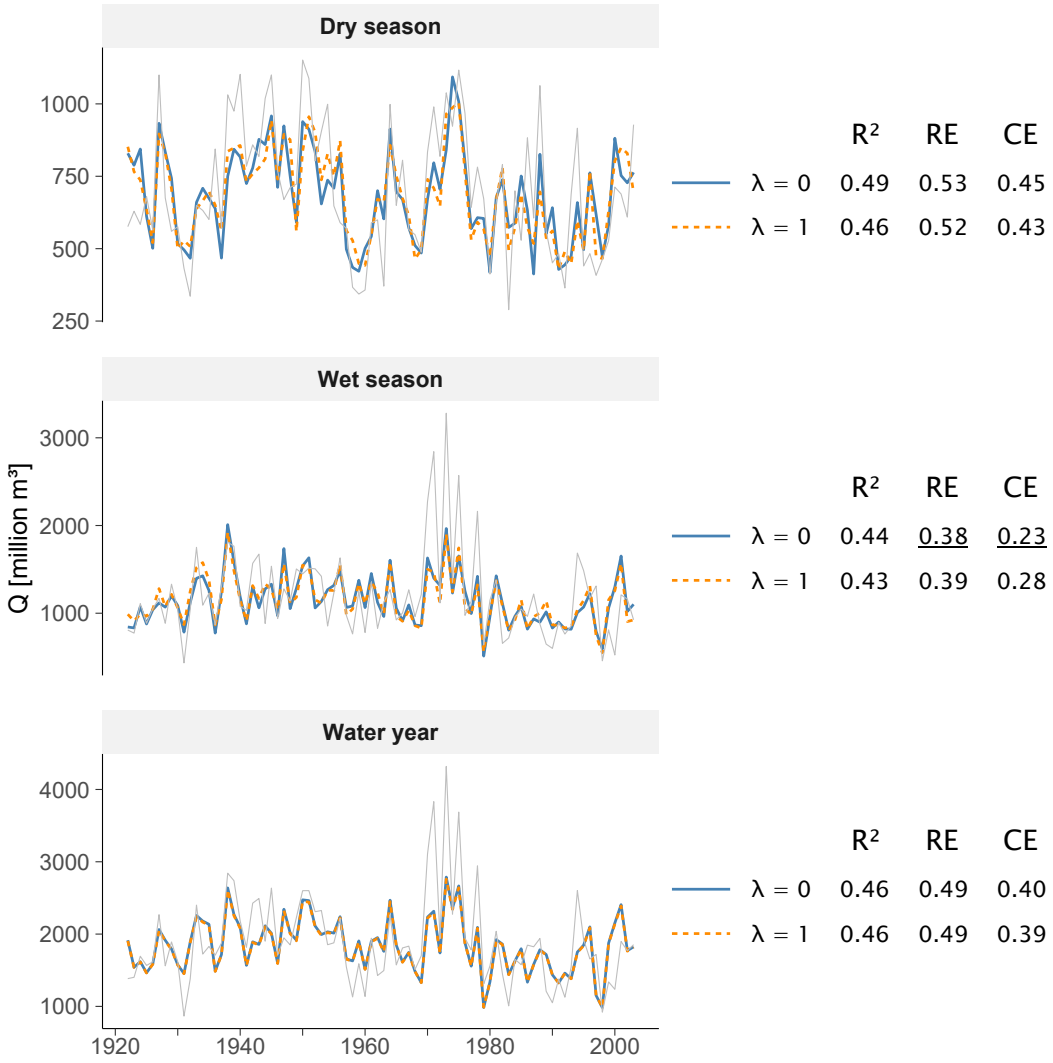

Figure 3.4. Reconstruction skills and time series for the instrumental period, produced using two models: without mass balance penalty ($\lambda = 0$) and with penalty ($\lambda = 1$) in the regression problem (Equation 3.14). Grey lines show naturalized observations. Underlined scores show where the reconstruction is not statistically skillful at $\alpha = 0.1$.

1 produces notably narrower distributions of RE and CE for the wet season (Figure 3.5). Consequently, Model 1's wet season reconstruction becomes statistically skillful. Overall, Model 1 is more robust.

To understand Model 1's robustness, let us recall the models' formulation. Model 0 reconstructs the dry season, wet season, and annual flows independently. Each reconstruction is geared towards its own target, and can become sensitive to training data—the wet season reconstruction does. Contrarily, Model 1 links all three reconstructions together via the penalty term (Equations 3.12–3.14). This link provides each reconstruction with additional information from the other two, thus preventing each reconstruction from overfitting to its own target. In our case, the wet season reconstruction

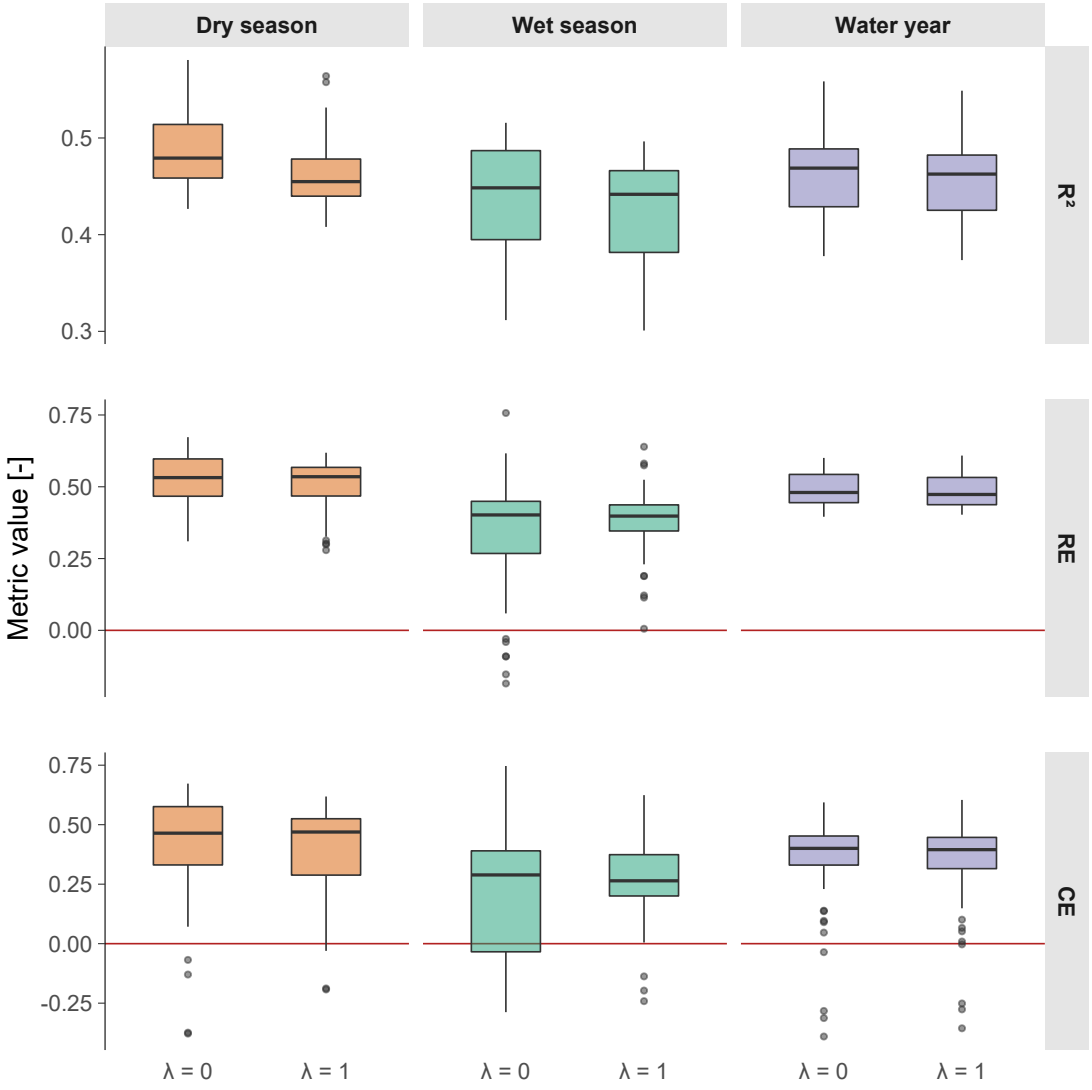

Figure 3.5. Distribution of skill scores.

benefited significantly from this additional information to become statistically skillful, with minimal trade-off from the other two reconstructions.

The selected input subsets by both models provide further insights into their similarities and differences. Both models use similar input subsets, with identical ones for the water year. However, Model 1 uses fewer inputs than does Model 0 for the wet and dry seasons (Figure 3.6). Therefore, the models behave similarly, but Model 1 is more parsimonious.

Surprisingly, the Mu Cang Chai and Mae Hong Son $\delta^{18}O$ chronologies were consistently selected by both models for all reconstructions, including the dry season reconstructions where we expected ring width chronologies to dominate, based on the observed tendencies of ring width to correlate more with wet season flow and $\delta^{18}O$ to

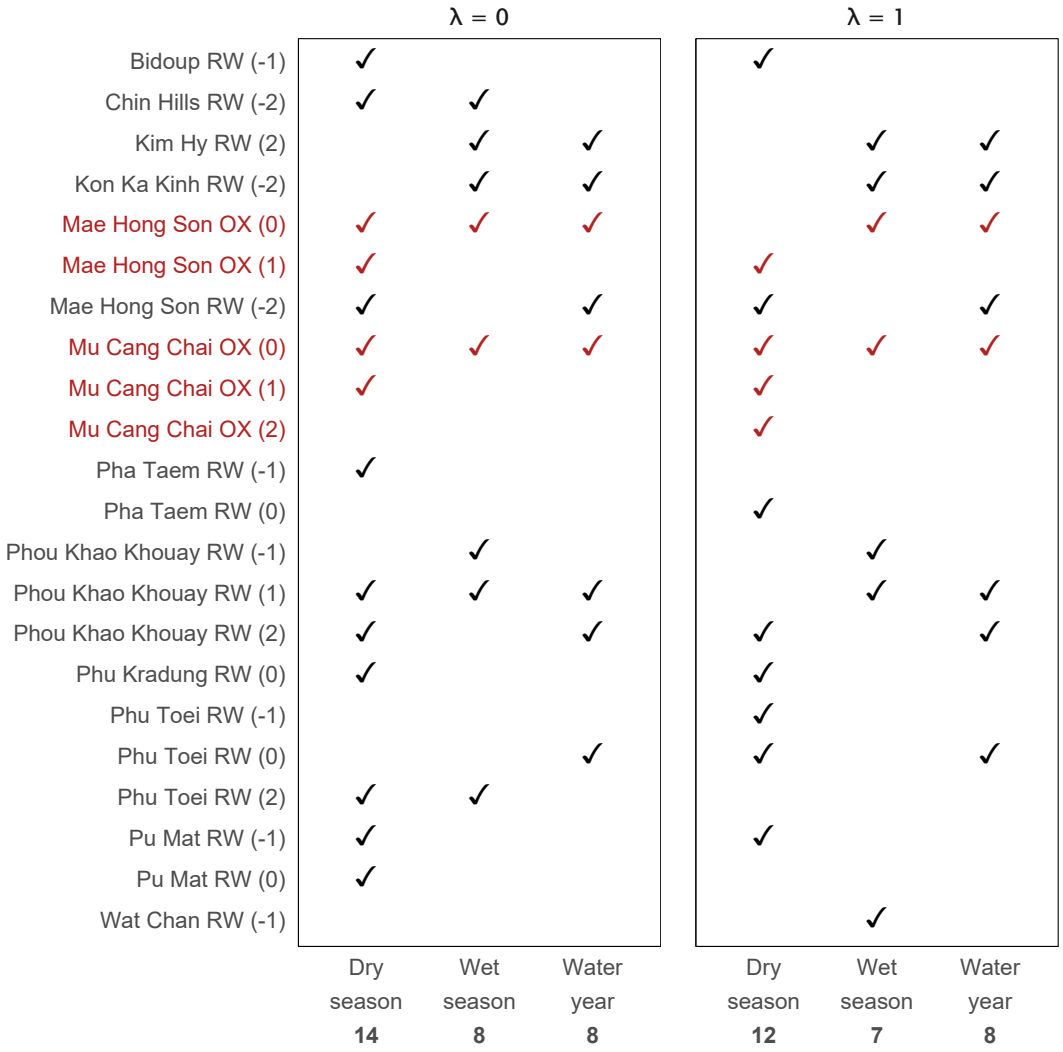

Figure 3.6. Selected inputs for each streamflow target (columns) in each model (panels). Only inputs that were selected at least once are shown. The stable oxygen isotope (OX) inputs are marked in dark red. Bold numbers at the bottom row of the x-axis are the number of selected inputs in each column. Both models selected similar input subsets. The number of inputs selected for each target by Model 1 is the same as or fewer than those selected by Model 0 for the same target.

correlate more with dry season flow. Upon closer examinations of the correlation analysis (Figure 3.3), the puzzle is solved. First, the $\delta^{18}O$ chronologies are intercorrelated C. Xu et al., 2019, and the input selection algorithm correctly selected the two sites that exhibit the strongest correlations. Second, 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.

Over the whole study horizon, we observe similar results to those of the instrumental period (Figure 3.7). The two models agree with each other. Model 1's adjustments are

generally small. The largest adjustments are in the wet season ($\pm$400 Mm^3); smaller but notable adjustments are seen in the dry season ($\pm$200 Mm^3), and minute adjustments ($<$ 90 Mm^3) are seen in the annual reconstruction. We will examine the effects of these adjustments in Section 3.4.2.

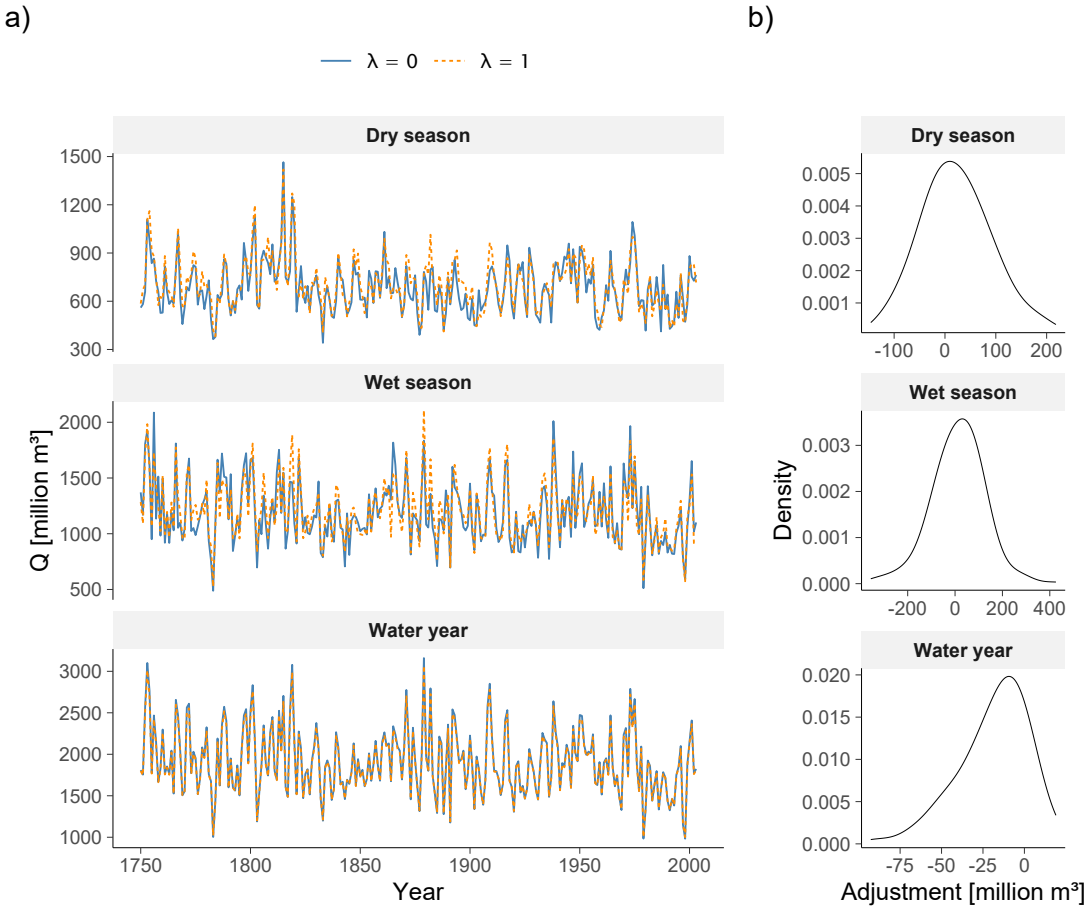

Figure 3.7. (a) Full reconstruction time series of both Model 0 ($\lambda = 0$) and Model 1 ($\lambda = 1$). (b) Distributions of the amounts of adjustment made by Model 1, that is, the estimates of Model 1 subtracted by those of Model 0.

The reconstructions provide some interesting insights into the inter- and intra-annual variability of the river (Figure 3.8). Between 1825–1855, 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 1790–1820, 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 8 of 254 years (Appendix B.6). Revealing events like this one is is of fundamental importance for the operations of Bhumibol Dam, which supports the activities of large irrigation districts.

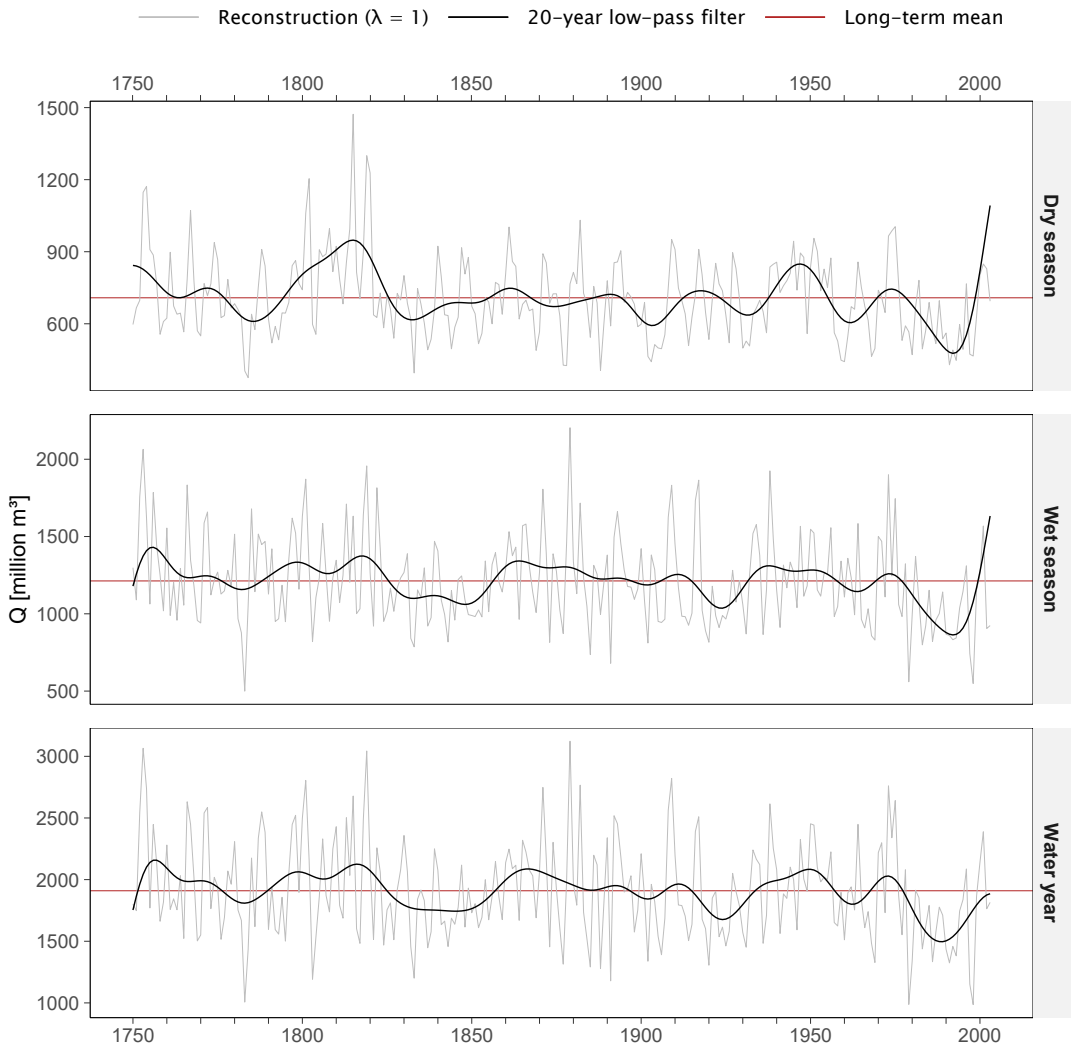

Figure 3.8. Full reconstruction results

3.4.2 Annual Mass Balance

For each model, we compare the total seasonal flow with the annual flow. To preserve the annual mass balance, these two time series should ideally be the same. However, for Model 0, large discrepancies are seen between the two time series (Figure 3.9a). For Model 1, the two time series agree with each other better. As each time series provides an estimate of the annual water budget, we are also interested in their distributions. We observe that the distributions produced by Model 0 are notably different from each other, but those produced by Model 1 are almost identical (Figure 3.9b). This implies that the distributions derived from Model 1 are more reliable. Using the same analysis, we show that Model 1 also produces more reliable distributions of the dry and wet season's water budget than does Model 0 (Figure B.12).

Next, for each model, we calculate the mass difference, ΔQ, between the total seasonal

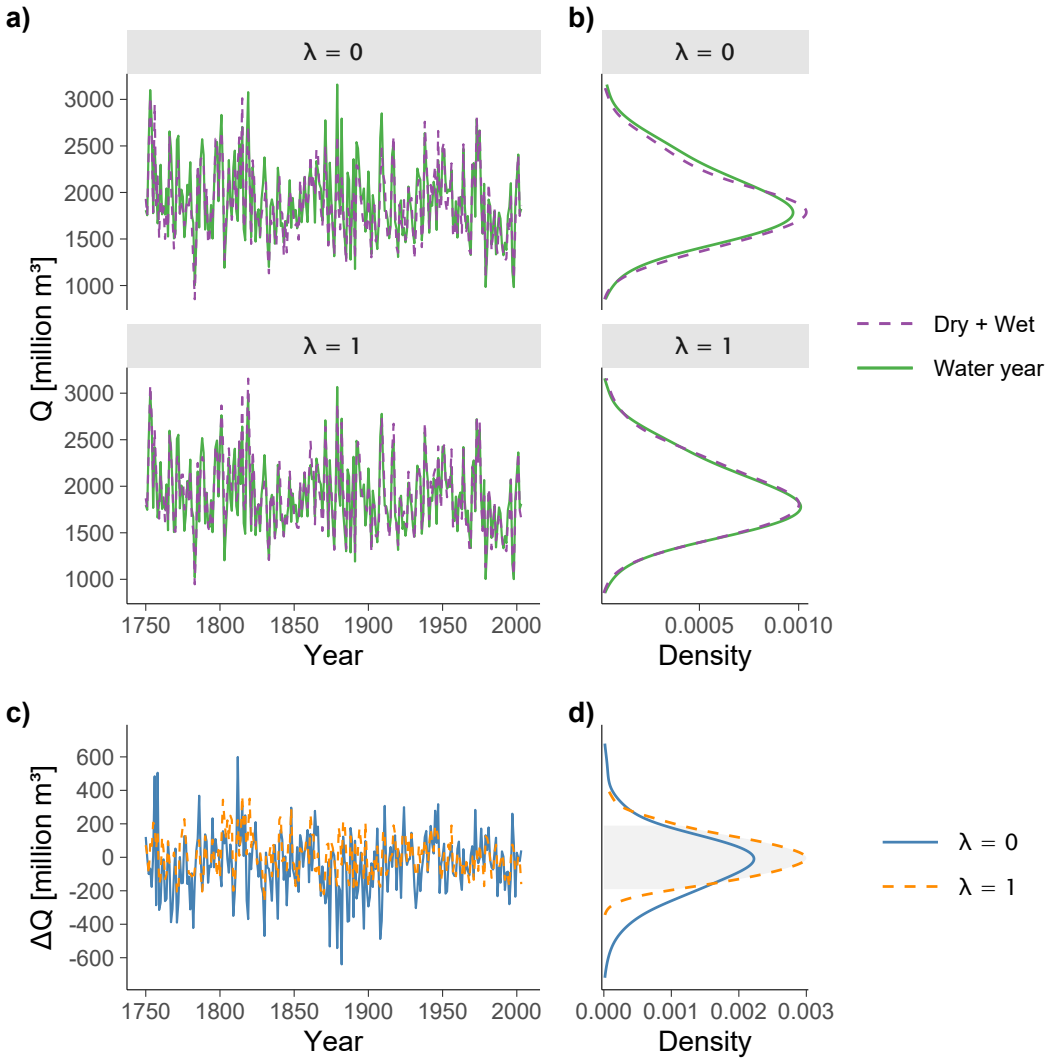

Figure 3.9. (a) Comparison between the total seasonal flow (TSF; Wet + Dry) and the annual flow (AF; Water year) for Model 0 ($\lambda = 0$) and Model 1 ($\lambda = 1$). (b) Distributions of the annual water budget, estimated by either the TSF or the AF for both models. (c) The differences, ΔQ, between the TSF and the AF. (d) Distributions of ΔQ. The shaded region denotes the ± 190 Mm^3 range, equivalent to $\pm 10\%$ of the mean annual flow.

flow and the annual flow, then examine its trajectory and distribution (Figures 3.9c and 3.9d). The mass difference for Model 0 ranges from -640 Mm^3 to 600 Mm^3, while that range for Model 1 is -270 Mm^3 to 370 Mm^3; a 50% reduction in range. Moreover, Model 0 yields a mass difference outside the interval ± 190 Mm^3 ($\pm 10\%$ of the mean annual flow; shaded region in Figure 3.9d) in 28% of the years. That figure for Model 1 is only 11%. By these metrics, Model 1 is twice better than Model 0 in terms of preserving mass balance.

3.5 Discussion and Conclusions

In Section 3.4.1, 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. More importantly, the adjustment significantly reduces the differences between the total seasonal flow and the annual flow (Section 3.4.2). Without the adjustment, the mass difference can be as large as 640 Mm^3, or about 30% of the mean annual flow. It amounts to 90% of the irrigation demand from the Ping River downstream of Bhumibol Reservoir (Divakar et al., 2011). Such a discrepancy may lead to water allocation disputes. With the adjustment, both the frequency and magnitude of discrepancies are reduced—this is crucial for water availability assessment, a major goal of dendrohydrology.

We also showed that the mass-balanced-adjusted regression produces reliable distributions of the seasonal and annual streamflow. These distributions can be used for probabilistic studies in water resources applications. For example, sub-annual stochastic time series can be generated from the distributions to be used in bottom-up vulnerability assessments of water systems (Pielke et al., 2012; Herman et al., 2016). In Appendix B.8 we illustrate one simple way to do so (by sampling from a bivariate distribution fitted to the seasonal reconstructions), but more advanced methods are available (e.g., Borgomeo et al., 2015a; Borgomeo et al., 2015b).

The framework that we proposed here can be reapplied and expanded in several ways. First, analysts adopting our framework have the choice to tune λ, depending on how important it is to preserve mass balance in their applications. In this case, a sensitivity analysis with respect to λ may be informative. Second, 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 difference between the total flow of the tributaries and the flow on the main stream. More proxies, such as maximum wood density and blue intensity, can also be incorporate to increase the amount of climate signal captured. Finally, the mass balance formulation is readily extendable to higher resolutions, e.g., quarterly or monthly, as long as the proxy network is sensitive enough to the higher resolution targets. (We will explore this in Chapter 5, Section 5.3.1.) These directions can help dendrohydrology realize its value in operational water management, an area where annual, unconstrained streamflow reconstructions have had limited success.

Chapter 4

Coherent Streamflow Variability in Monsoon Asia over the Past Eight Centuries—Links to Oceanic Drivers

Publication

Nguyen, H. T. T., Turner, S. W. D., Buckley, B. M., & Galelli, S. (2020b). Coherent stream flow variability in Monsoon Asia over the past eight centuries — links to oceanic drivers. *Water Resources Research*, *e2020WR027*. https://doi.org/10.1029/2020WR027883

4.1 Introduction

Of the world's 30 biggest rivers, ten are located within Monsoon Asia, and two others originate from this region (Figure 4.1). These river basins are home to 1.7 billion people (Best, 2019). With high population densities, even smaller basins support the livelihood of millions—e.g., Chao Phraya (Thailand): 25 million, Angat (the Philippines): 13 million, and Citarum (Indonesia): 10 million (Nguyen & Galelli, 2018; Libisch-Lehner et al., 2019; D'Arrigo et al., 2011a). River discharge, or *streamflow*, provides water for domestic and industrial uses, irrigation, and hydropower. It sustains aquatic life (including fish yield), carries sediment and nutrients, helps prevent salinization of river deltas, and facilitates navigation. Streamflow is an important link in both the water-energy-food nexus and the ecological cycle. To manage this resource, we need a good understanding of hydrologic variability. Such understanding is often derived from streamflow measurements; however, these instrumental data span typically only a few decades, too short to capture long-term variability and changes in streamflow.

When compared against instrumental data, longer streamflow records reconstructed from climate proxies—such as tree rings—often reveal striking insights. A reconstructed pre-dam variability of the Yellow River (Li et al., 2019) shows that streamflow in 1968–2010 was only half of what should have been; in other words, human withdrawals for agriculture, industry, and municipalities reduced streamflow by half! A reconstruction of the Citarum River (Indonesia) (D'Arrigo et al., 2011a) shows that the period 1963–2006 contained an increasing trend of low flow years but no trend in high flow years, compared with the previous three centuries. This finding suggests that

10 million inhabitants of Jakarta may be facing higher drought risks than what is perceived from the instrumental record. The Mongolian "Breadbasket", an agricultural region in north-central Mongolia (Pederson et al., 2013), experienced an unusually wet twentieth-century, and the recent dry epoch is not rare in the last four centuries (Davi et al., 2006; Pederson et al., 2013; Davi et al., 2013). Consequently, agricultural planning cannot take the twentieth century to be the norm, lest history repeats the lesson of the Colorado River Basin: observations over abnormally wet years (Stockton & Jacoby, 1976; Woodhouse et al., 2006; Robeson et al., 2020) led to water rights over-allocation, and the Colorado no longer reaches the Pacific Ocean.

The case of the Colorado River demonstrates that streamflow reconstructions can improve our understanding of water resources availability. Furthermore, with longer streamflow records, low frequency variations of streamflow can be revealed, the frequency and magnitude of floods and droughts can be better quantified, and the risks associated with these natural disasters can be better assessed. These benefits have been demonstrated in Australia (Allen et al., 2017; Tozer et al., 2018), the United States (DeRose et al., 2015; Stagge et al., 2018), Canada (Hart et al., 2010; Sauchyn et al., 2015) and other countries (Lara et al., 2015; Güner et al., 2017). Streamflow reconstructions have also been used to generate stochastic time series for water management applications (Prairie et al., 2008; Sauchyn & Ilich, 2017). These benefits, if realized in Monsoon Asia, can improve the lives of many people, given the dense populations that inhabit river basins in this region.

Compelling evidence calls for more streamflow reconstructions in Monsoon Asia. Tremendous efforts, particularly in the last four years (Figure C.1), have partly addressed this need, but the hydrological knowledge gained was limited to individual catchments, more than half of which are in China (Table C.1). A regional, synthesized understanding is lacking. Here, we produce the first large-scale streamflow reconstruction for Monsoon Asia, covering 62 stations in 16 countries, unraveling eight centuries of annual streamflow variability. To achieve this task, we develop a novel automated framework with three main components: (1) a climate-informed proxy selection procedure, (2) a dynamic state-space reconstruction model, and (3) a rigorous cross-validation routine for parameter tuning to achieve optimal skills. We also use the Monsoon Asia Drought Atlas version 2 as the paleoclimatic proxy instead of a tree ring network, as the former offers computational advantages (supported with strong physical and statistical foundations) for this large-scale reconstruction. With this work, 58 stations are reconstructed for the first time while the other four (Citarum, Yeruu, Ping, and Indus Rivers) are extended back in time compared to previous works (D'Arrigo et al., 2011a; Pederson et al., 2013; Nguyen & Galelli, 2018; Rao et al., 2018). This data set allows us to assess both local historical water availability and regional streamflow patterns, revealing the spatial coherence of streamflow and its links to the oceans. This understanding may improve the management of river basins and other water-dependent resources.

4.2 Data

4.2.1 Streamflow Data

Our reconstruction target is the mean annual flow, and we used the calendar year (January to December) as there is not a common water year across Monsoon Asia (Knoben et al., 2018). We obtained streamflow data from the Global Streamflow Indices and Metadata Archive (GSIM) (Do et al., 2018; Gudmundsson et al., 2018), using stations having at least 41 years of data, and with less than 3% missing daily values. We also received streamflow data from our colleagues for some countries where public streamflow records are not available (see Appendix C.2). Small catchments may be influenced by local conditions more than by broad climate inputs that are captured in the regional paleoclimate proxies (Strange et al., 2019). Therefore, we used only stations where the mean annual flow over the whole time series is at least 50 m^3/s; this threshold is heuristic, and somewhat arbitrary. Details of this initial selection step are provided in Appendix C.2 and in the code repository for this paper (ntthung.github.com/paleo-asia, DOI: 10.5281/zenodo.3818116.)

Many stations in our collection have upstream reservoirs that may interfere with the proxy-streamflow relationship. This interference is stronger for seasonal streamflow than annual streamflow: reservoirs transfer water from the wet season to the dry season, but not all reservoirs retain water from year to year. Reservoirs that are filled and emptied within a year do not change the annual water budget downstream. To minimize reservoir interference, we reconstructed annual streamflow, and we removed stations that have upstream retention times longer than a year. We identified upstream reservoirs by overlaying the Global Reservoirs and Dams (GRanD) data (Lehner et al., 2011) on the river network (Lehner & Grill, 2013; Barbarossa et al., 2018). The upstream retention time was calculated as the total upstream reservoir capacity (million m^3) divided by the mean annual flow volume (million m^3/year). For stations having over-year reservoirs constructed towards the end of their records, we also truncated the corresponding years, keeping only the streamflow data before dam construction.

Our collection and quality control effort resulted in an annual streamflow data set of 62 stations in 16 countries. Our records span across Monsoon Asia, covering the following sub-regions: Central Asia (CA), East Asia (EA), eastern China (CN), West Asia (WA), Southeast Asia (SEA), and South Asia (SA). The stations' locations and upstream retention times (for those having upstream reservoirs) are shown in Figure 4.1.

4.2.2 Proxy Data

4.2.2.1 The Monsoon Asia Drought Atlas version 2

Our paleoclimate proxy is the Monsoon Asia Drought Atlas version 2 (MADA v2) (E. R. Cook, 2015), built upon the original MADA of E. R. Cook et al. (2010a). The MADA is a gridded data set of the Palmer Drought Severity Index (PDSI) (W. C. Palmer, 1965) over the Asian monsoon region; each grid cell contains an annual time series of the mean June-July-August PDSI, reconstructed from tree rings, and calibrated with the instrumental data set of Dai et al. (2004). The MADA proves to be a reliable long-term record of monsoon strength, having revealed the spatiotemporal extents of the

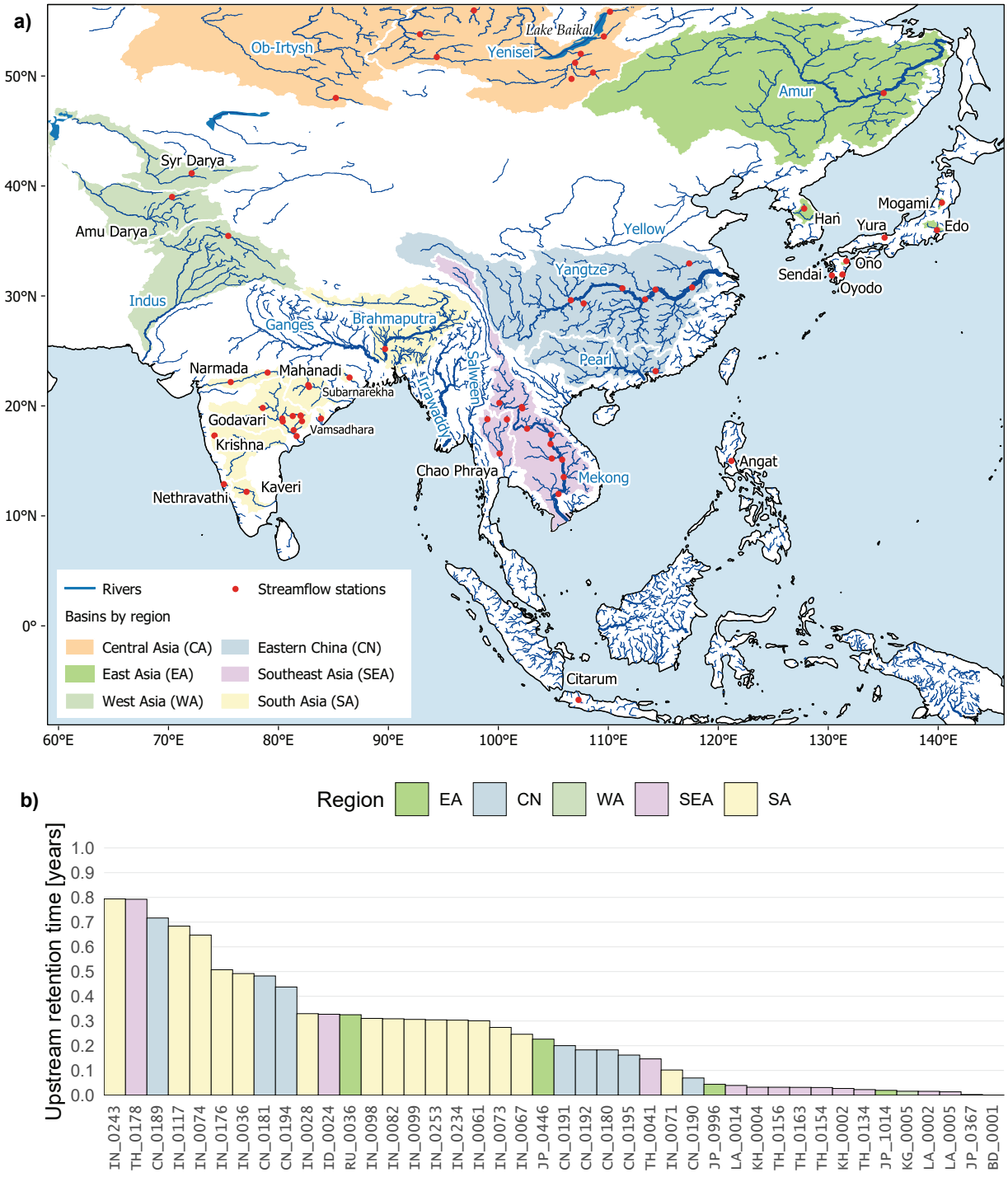

Figure 4.1. a) The Monsoon Asia region (E. R. Cook et al., 2010a); river basins involved in this study are highlighted by sub-region, rivers belonging to the world's 30 biggest (Best, 2019) shown with blue names. b) Upstream retention time of the 42 stations that have upstream reservoirs. The bar colours denote the regions according to a). The first two letters of each station's code indicates the country it is in.

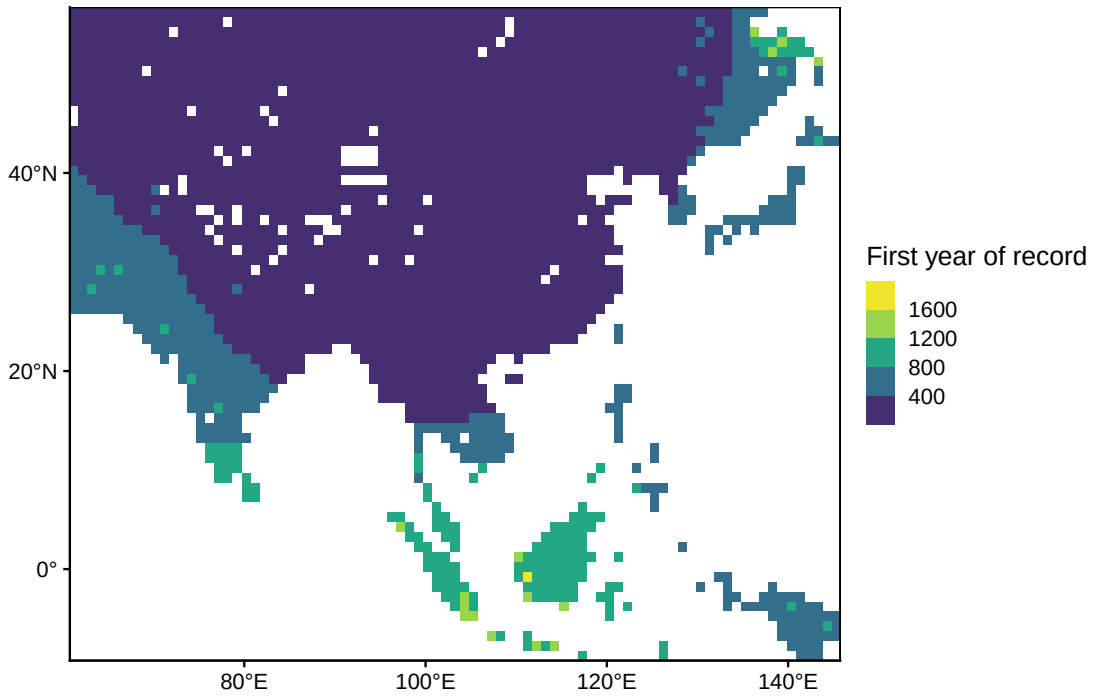

Figure 4.2. First year of record for each MADA grid point. 2716/2732 grid points start at or before 1200. The remaining grid points are not used.

four Asian megadroughts in the last millennium, and linking variations in monsoon strength to sea surface temperature patterns. MADA v2 improves over its predecessor by incorporating more tree ring chronologies (453 versus 327), and targeting the self-calibrating PDSI (scPDSI), which addresses several limitations of the standard PDSI (Wells et al., 2004; van der Schrier et al., 2013). We use the MADA v2 portion between 1200–2012 as this is the common period of most grid points in the atlas (Figure 4.2), and is also the stable portion with sufficient number of tree ring chronologies in the source tree ring network.

4.2.2.2 Computational advantages of using the MADA, and caveats

The construction of the MADA can be thought of as a transformation from the tree ring network, irregular in both space and time, to a regular grid with homogeneous temporal coverage—analogous to transforming meteorological station data to gridded temperature and precipitation products. This transformation brings several advantages to reconstructing streamflow using the MADA, compared to using the underlying tree ring network.

First, in a typical reconstruction study, one must detrend and standardize the tree ring data to remove non-climate signals (cf E. R. Cook & Kairiukstis, 1990). For a large scale study like ours, such a task is complex. Instead, we can leverage the effort that has been devoted to detrending and standardizing the chronologies in making the MADA, and use the MADA as proxy, having built the physical and statistical foundations to do so (Section 1.2).

Second, the tree ring sites often cluster, with vast empty space between clusters (see e.g. E. R. Cook et al. (2010a), Figure 1). When taking a subset of them for reconstruction at a station, there can be cases where none or very few sites are within a search radius. The MADA helps "bridging" the space, bringing climate signals from further-away tree sites to grid points nearer to the station. The high resolution grid ($1^\circ \times 1^\circ$ for version 2) makes automated grid point selection easier. (The automated grid point selection procedure is described in Section 4.3.1.)

Third, when reconstructing streamflow from tree rings, nested models are often necessary because tree ring chronologies have different time spans. One starts with the shortest nest, using the common time span of all chronologies to build a model, then dropping the shortest chronology to build a second model with longer time span but less explained variance than the first, and repeating the process, dropping more chronologies to achieve longer time spans until the final nest with the longest time span, but with the lowest explained variance. The nests' outputs are then corrected for their variance and averaged to obtain the final reconstruction (see e.g., D'Arrigo et al., 2011a). This nesting step was carried out for the MADA, such that most grid points have the same time span (Figure 4.2). This lets us use a single common period (1200-2012), and eliminates our need to build nested models back in time. This is particularly desirable for our dynamic state-space reconstruction model, as averaging the nests breaks the link between the catchment state and streamflow. (The reconstruction model is described in Section 4.3.2.)

The computational advantages of using the MADA are thus threefold: (1) no detrending and standardization, (2) easier grid point selection, and (3) no nesting. However, these come with some costs, the most important of which is uncertainty. When reconstructing streamflow from the MADA, we treat the MADA (i.e., the model input) as constant. But in fact, the MADA is a regression product and has its own uncertainty. Furthermore, this uncertainty increases back in time as the number of available chronologies decreases. One way to quantify the uncertainty is by bootstrapping: streamflow reconstructions can be built using bootstrap replicates of the MADA, and the range of the bootstrap ensemble indicates the uncertainty of the reconstruction. An appropriate bootstrapping scheme must be considered, given that the uncertainty is nonstationary, and that dimensionality is a challenge: the MADA has 813 years $\times$ 2716 grid points. In this regard, the added benefit of our reconstruction framework is that it runs for each station individually (see Section 4.3), so one need not reconstruct the whole network in order to quantify uncertainties at specific stations of interest.

As a gridded regression product, the MADA smooths out local variability. This can be alleviated by carefully selecting and processing the grid points to retain as much variance as possible (Section 4.3.1), and by using sufficiently large catchments (Section 4.2.1).

Finally, we note that the computational advantages we described here are only applicable to large-scale studies, where an automated framework is needed. For individual sites, we urge researchers to consider all available proxies, rather than being attracted by the convenience offered by the drought atlases.

4.3 Point-by-Point, Climate-informed, Dynamic Streamflow Reconstruction

When reconstructing a climate field, such as a PDSI grid or a streamflow station network, it is desirable to preserve the field covariance structure. However, building a large-scale spatial regression model is challenging. Instead, one can reconstruct each point in the field independently, and rely on the proxy network to capture the spatial patterns. This is the premise of the Point-by-Point Regression (PPR) method (E. R. Cook et al., 1999), which has been used to reconstruct drought atlases of Europe (E. R. Cook et al., 2015; E. R. Cook et al., 2020), the Americas (E. R. Cook et al., 1999; Stahle et al., 2016; Morales et al., 2020), Oceania (J. G. Palmer et al., 2015), and Asia (E. R. Cook et al., 2010a). These drought atlases demonstrate that PPR captures well the spatial patterns of climate variability (see e.g. E. R. Cook et al. (1999), Figures 8 and 9). Like these drought atlases, our streamflow network covers a large spatial domain with varying climates; therefore, we adopted the PPR principle, and reconstructed our stations individually. While some aspects of our reconstruction framework followed the PPR procedure, we have innovated many steps of the workflow.

Overall, the framework involves three main stages: (1) input selection (Section 4.3.1), (2) model calibration (Section 4.3.2), and (3) cross-validation (Section 4.3.3). In Stage 1, we selected a subregion of the MADA that is hydroclimatically similar to the streamflow station of interest, and extracted from this subregion a parsimonious subset of principal components, using weighted Principal Component Analysis (PCA). This stage involves two tuning parameters: the hydroclimate similarity threshold, and the PCA weight. For each combination of these parameters, we calibrated a reconstruction model in Stage 2, thus producing an ensemble of models. Finally, in Stage 3, we cross-validated the models to choose the best one, and used that for the final reconstruction.

4.3.1 Climate-informed Input Selection

A regional paleoclimate proxy record, such as the MADA or its underlying tree ring network, is rich with information, but not all of such information is relevant to the streamflow target. A proper input selection is necessary to filter noise and retain only the most relevant signal. A common way is to use proxy sites within a search radius; and PPR does the same. But, given that geographical proximity does not necessarily imply hydroclimatic similarity, we selected our proxies (MADA grid points) by hydroclimatic similarity directly. The hydroclimate at location i (a MADA grid point or a streamflow station) is characterized by three indices: aridity a_i, moisture seasonality s_i, and snow fraction f_i, following Knoben, Woods, and Freer, who proposed this hydroclimate characterization and calculated the indices for a global $0.5^\circ \times 0.5^\circ$ grid (Knoben et al., 2018). The hydroclimatic similarity between two locations i and j is then defined as their Euclidean distance in the hydroclimate space. This distance is termed the KWF distance and its mathematical definition is

$$d_{KWF}(i,j) = \sqrt{(a_i - a_j)^2 + (s_i - s_j)^2 + (f_i - f_j)^2}. \tag{4.1}$$

By calculating the KWF distance between each MADA grid point and each streamflow station, we can screen out MADA grid points that are geographically close to the station of interest but hydroclimatically different—a climate-informed grid point selection scheme. Whereas previous PPR implementations varied the search radius, we fixed the radius to 2,500 km—the scale of regional weather systems (Boers et al., 2019)—and varied the KWF distance between 0.1 and 0.3 in 0.05 increments. For reference, the maximum KWF distance between any two points in Monsoon Asia is 1.424. Each KWF distance yielded a search region encompassing a set of MADA grid points surrounding the streamflow station of interest. In our search regions, PDSI often correlates significantly and positively with streamflow (Figure 4.3); indeed hydroclimatic similarity is a physical basis for correlation.

Next, we performed weighted PCA to remove multicollinearity among the MADA grid points. Following PPR, we weighted each grid point by its correlation with the target streamflow, using Equation 4.2:

$$z_i = g_i r_i^p. \tag{4.2}$$

Here, g_i is grid point i's scPDSI time series, r_i the correlation between g_i and the target streamflow, p the weight exponent, and z_i the weighted version of g_i. We used $p = 0, 0.5, 2/3, 1, 1.5$, and 2, the same as those used by E. R. Cook et al. (2010a). We then performed PCA on z_i's, and retained only those principal components (PCs) having eigenvalue at least 1.0 (Hidalgo et al., 2000). From the retained PCs (typically about 20–40 per station), we selected a parsimonious subset that is most relevant to the streamflow target using the VSURF (Variable Selection Using Random Forest) algorithm (Genuer et al., 2010). So, for each combination of KWF distance and PCA weight, we arrived at a subset of PCs for reconstruction. Each streamflow station has an ensemble of 30 such subsets, the best of which was identified using cross-validation (Section 4.3.3) and used for the final reconstruction.

4.3.2 Linear Dynamical System

Having obtained the climatic inputs, the next step was to model the relationship between these inputs and the catchment output (streamflow). Here, this relationship was not modeled with linear regression (as with original PPR, and as typical with previous reconstruction studies), but as a linear dynamical system (LDS), following Equations 4.3 and 4.4:

$$x_{t+1} = Ax_t + Bu_t + w_t \tag{4.3}$$

$$y_t = Cx_t + Du_t + v_t \tag{4.4}$$

where t is the time step (year), y the catchment output (streamflow), u the climatic input (an ensemble member from the climate-informed grid point selection), w and v white noise, and x the hidden system state, which can be interpreted as the catchment's flow regime, i.e, wet or dry (Nguyen & Galelli, 2018). By modeling the flow regime and its transition, the LDS model accounts for both regime shifts (Turner & Galelli, 2016) and catchment memory (Pelletier & Turcotte, 1997). These behaviors are not modeled in linear regression.

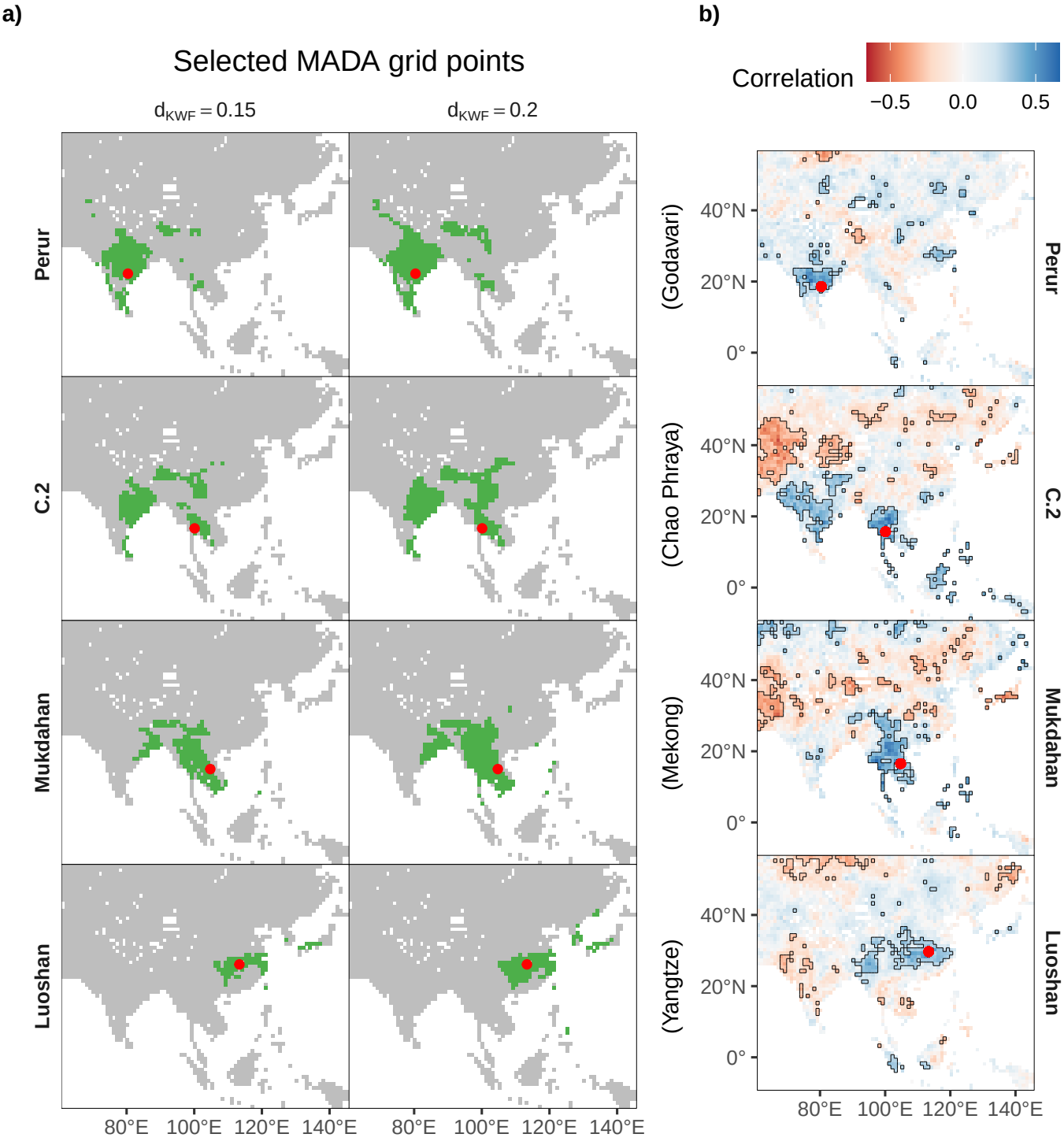

Figure 4.3. a) Examples of the climate-informed grid point selection: selected MADA grid points (green) based on two KWF distances (columns) at four stations (rows). b) Correlations between streamflow at the same four stations and the MADA, significant correlations ($\alpha = 0.05$) enclosed in black boundaries. The selection regions in (a) generally have significant positive correlation with streamflow. The areas with significant negative correlations need further investigation.

The LDS model assumes that the initial state and the noise processes are normally distributed:

$$w_t \sim \mathcal{N}(0, Q) \tag{4.5}$$
$$v_t \sim \mathcal{N}(0, R) \tag{4.6}$$
$$x_1 \sim \mathcal{N}(\mu_1, V_1). \tag{4.7}$$

It follows that the catchment state and output are also normally distributed. But some of our streamflow records are skewed. These were log-transformed to reduce skewness (Appendix C.3).

The LDS model is trained using a variant of the Expectation-Maximization algorithm. In the E-step, we fix the model parameters and learn the hidden state. In the M-step, we fix the hidden state and learn the model parameters. Iterations are repeated between the E- and M-steps until convergence. The reconstruction algorithm is implemented in the R package *ldsr* (Nguyen, 2020).

4.3.3 Cross-validation

Consistent with the literature, we assessed reconstruction performance using the metrics Reduction of Error (RE) and Nash-Sutcliffe Coefficient of Efficiency (CE or NSE) (Nash & Sutcliffe, 1970; Fritts, 1976). Mathematically,

$$RE = 1 - \frac{\sum_{t \in \mathcal{V}} (Q_t - \hat{Q}_t)^2}{\sum_{t \in \mathcal{V}} (Q_t - \bar{Q}_c)^2} \tag{4.8}$$

$$CE = 1 - \frac{\sum_{t \in \mathcal{V}} (Q_t - \hat{Q}_t)^2}{\sum_{t \in \mathcal{V}} (Q_t - \bar{Q}_v)^2} \tag{4.9}$$

where t is the time step, $\mathcal{V}$ the validation set, Q the observed streamflow, $\hat{Q}$ the reconstructed streamflow, $\bar{Q}_c$ the calibration period mean, and $\bar{Q}_v$ the verification period mean.

Both RE and CE are based on squared error; they can be sensitive to outliers, especially the CE. To address this limitation, Gupta et al. (2009) proposed another metric, which assesses a model output based on its correlation with observation, as well as its bias and variability (Equation 4.10):

$$KGE = 1 - \sqrt{(\rho - 1)^2 + \left(\frac{\hat{\mu}}{\mu} - 1\right)^2 + \left(\frac{\hat{\sigma}}{\sigma} - 1\right)^2}. \tag{4.10}$$

Here, ρ is the correlation between model output and observation, $\hat{\mu}$ and μ the modeled and observed mean of the streamflow time series, and $\hat{\sigma}$ and σ the modeled and observed standard deviation of the streamflow time series. This metric is now known as the Kling-Gupta Efficiency (KGE). The KGE complements RE and CE, and we included the KGE in model assessment.

Conventionally, reconstruction skills are often calculated in a split-sample (i.e., two-fold) cross-validation scheme: the model is calibrated with the first half of the data and validated with the second half, then calibrated with the second half and validated with the first half (see e.g., D'Arrigo et al., 2011a). The contiguous halves aim to test a model's ability to capture a regime shift (Briffa et al., 1988). Unfortunately, this scheme is not practical for many stations in our record, where it would leave us only 20–25 data points for calibration (Figure S2). In addition, a two-fold cross-validation scheme provides only two point estimates for each skill score, and they may be notably different (for example, D'Arrigo et al. (2011a) reported CE values of 0.21 and 0.73 for the two folds.) As a result, the mean skill score may not be robust. A number of recent works have instead used the leave-k-out cross-validation scheme (e.g., Gallant & Gergis, 2011; Ho et al., 2016; Li et al., 2019). In this scheme, a random chunk of k data points is withheld for validation while the model is calibrated with the remaining data points, then calibration and validation are repeated over as many as 100 chunks of k. This scheme provides a more robust estimate of the mean skill score, but it may not correctly assess the model's ability to capture a regime shift, because the withheld points are not contiguous like in the split-sample scheme.

We sought a balanced approach. In each cross-validation run for each station, we withheld a *contiguous* chunk of 25% of the data points for validation and trained the model on the remaining 75%. This way, we maintain the goal of the split-sample scheme while still having enough data for calibration and getting distributions of skill scores, which yield a reasonably robust mean skill estimate for each metric. Having distributions of skill scores has another benefit: we can now make probabilistic statements about skill. For example, we can calculate the probability that $CE < 0$, and if that probability is less than a threshold α, say 0.1, then we consider the reconstruction statistically skillful with respect to CE at $\alpha = 0.1$. While not doing formal statistical tests, we can make analogous statements about the significance of the skills scores.

When the hold-out chunks are contiguous, there are not as many chunks as when they are random, so we repeated the procedure 30 times instead of 100, and calculated the mean RE, CE and KGE over these 30 runs. When calculating the mean scores, we used the Tukey's biweight robust mean (Mosteller & Tukey, 1977) instead of the arithmetic mean, to limit the effect of outliers. The robust mean is commonly used by dendrochronologists to derive the mean chronology from tree ring samples (E. R. Cook & Kairiukstis, 1990), and we have extended its use here. After cross-validating all ensemble members (Section 4.3.1), we selected one member for each station based on the robust mean CE and KGE (RE is similar to CE and is omitted). The ideal score for both CE and KGE is 1; therefore, we calculated for each ensemble member the Euclidean distance between the tuple (CE, KGE) and the point (1, 1). For each station, the ensemble member nearest to the ideal score was used for the final reconstruction.

4.4 Results and Discussion

4.4.1 Reconstruction Skills

Reduction of Error (RE) is positive at all stations (Figure 4.4a and b); Coefficient of Efficiency (CE) is positive at all but one: Kachora in the Indus (Pakistan), where CE ≈ -0.06 (Figures 4.4c and d). At $\alpha = 0.1$, 30 stations are statistically skillful with respect to RE, and 23 are with CE (Figure S9). Comparing the histograms of RE and CE (Figures 4.4b and d), we observe that CE is slightly lower—this is expected as CE is a more stringent metric than RE (E. R. Cook & Kairiukstis, 1990). Much lower CE than RE implies overfitting; we do not observe that here.

When using the Kling-Gupta Efficiency (KGE), if one wishes to benchmark a model against the verification period mean (as is with the CE), the threshold value is $1 - \sqrt{2} \approx -0.41$, i.e, KGE > -0.41 is analogous to CE > 0 (Knoben et al., 2019). Our KGE ranges from 0.22 to 0.68 (Figure 4.4e and f), far higher than the threshold. Furthermore, all 62 stations are statistically skillful with respect to KGE at $\alpha = 0.1$ (Figure 4.5). These results indicate that our reconstruction model performs well in terms of key characteristics: correlation, bias, and variability.

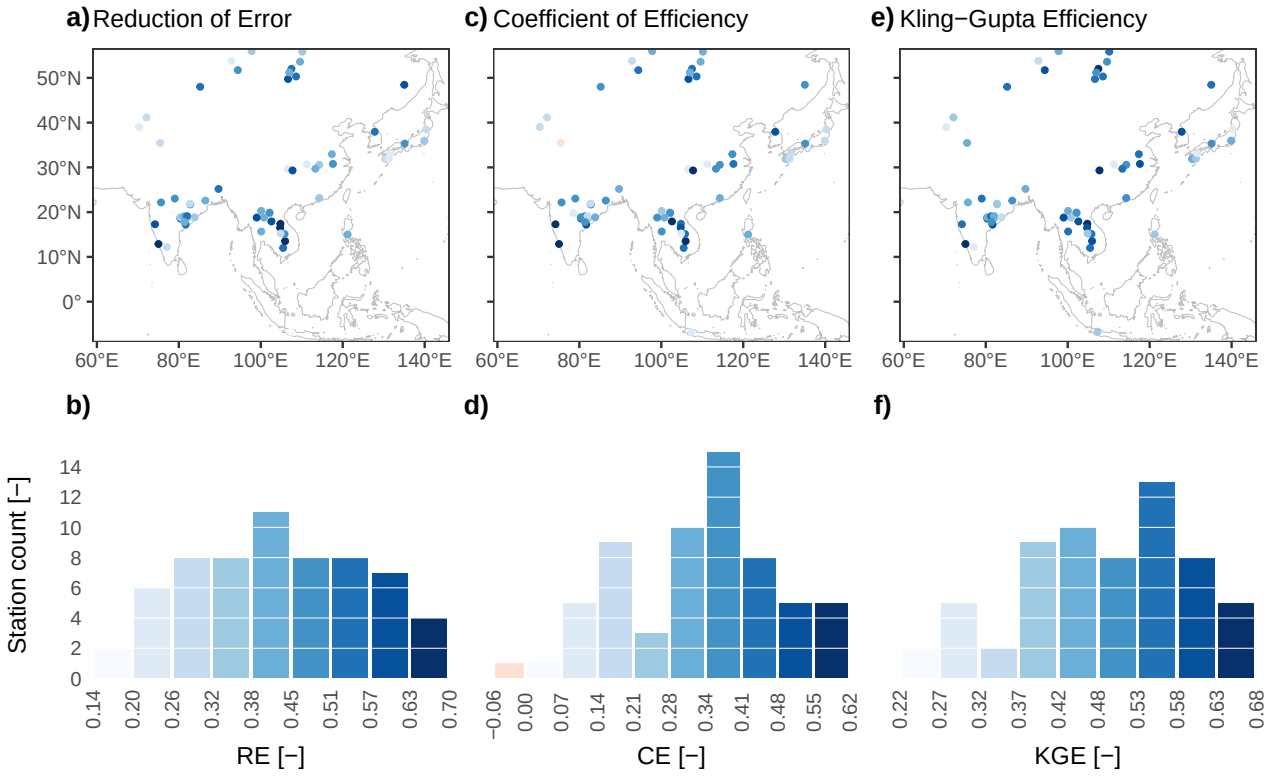

Figure 4.4. Distribution of model performance scores. Panels a, c, and e show the scores of each station following the color legends encoded with the histograms in panels b, d, and f.

All three metrics have similar spatial distributions (Figure 4.4a, c, and e). As expected,

lower skills are seen in most of Central Asia, Japan, and West Asia, which lie outside the core monsoon area. An exception is the upper reach of the Selenge River, upstream of Lake Baikal, where model skill is high, owing to high quality tree ring records from Mongolia (Davi et al., 2006; Pederson et al., 2013; Davi et al., 2013; Pederson et al., 2014). In all other regions, model skill is homogeneous. The consistent performance of our model suggests that the MADA is a good proxy for streamflow reconstruction in Asia, and our climate-informed dynamic reconstruction is skillful. More validation exercises (Appendix C.4) further support the reliability of the reconstruction.

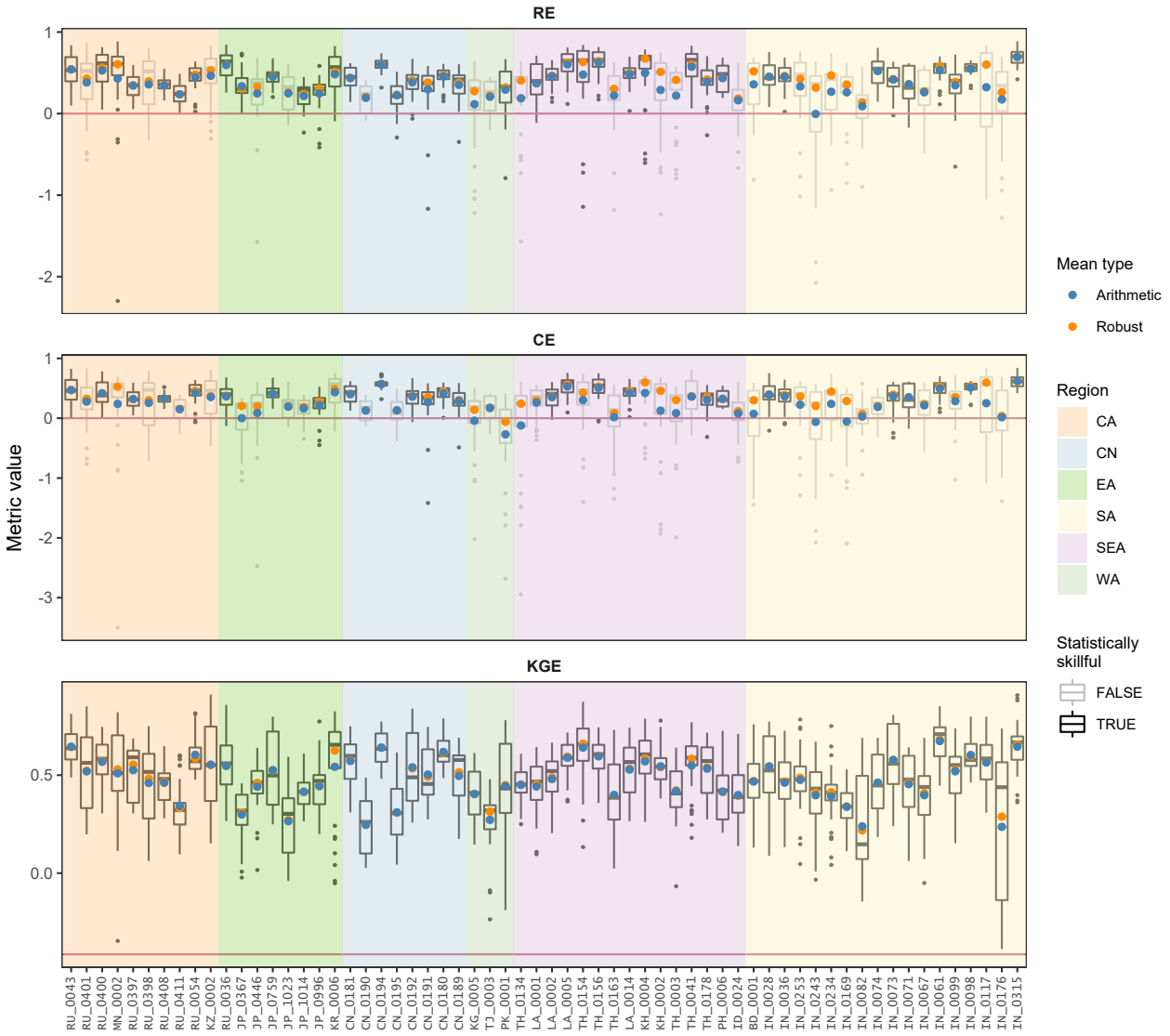

Figure 4.5. Distribution of performance scores over cross-validation runs for each station.

4.4.2 Spatiotemporal Variability of Monsoon Asia's Streamflow

Having obtained reliable skill scores, we now present eight centuries of spatiotemporal streamflow variability in Monsoon Asia, in terms of standardized streamflow (z-score of mean annual flow) (Figure 4.6a). This reconstructed history captures the riparian footprint of major historical events—large volcanic eruptions and megadroughts (Figure 4.6b). We first discuss the impact of the three largest eruptions of the past eight centuries (Sigl et al., 2015): Samalas (1257) (Lavigne et al., 2013), Kuwae (1452-53) (Gao et al., 2006), and Tambora (1815) (Stothers, 1984).

Assuming that Kuwae erupted in 1452 (consistent with tree ring records, see e.g. Briffa et al. (1998)), these three eruptions saw similar streamflow patterns (Figure 4.6b, panels 1, 4, and 8). In the eruption year t ($t = 1257, 1452, 1815$), large positive streamflow anomalies were observed in Southeast and East Asia. The magnitude of the positive anomalies were largest with Samalas, followed by Kuwae, and then Tambora. The global radiative forcings of the Samalas, Kuwae, and Tambora events are -32.8, -20.5, and -17.1 W/m^2, respectively (Sigl et al., 2015). Thus, we observe a correspondence between the magnitude of positive streamflow anomalies and the magnitude of radiative forcings. This correspondence is also seen clearly from the distributions of streamflow anomalies in the three events (Figure 4.7). These results suggest an influence of volcanic eruptions on streamflow in Southeast and East Asia.

Unlike East and Southeast Asia, South Asia's streamflow remained around the normal level in years t and $t+1$ in all three eruptions, suggesting little volcanic influence. More patterns were also observed: mixed wet and dry conditions in Central Asia, and normal to wet conditions in eastern China and West Asia (Figure 4.7a). Thus, the influence of volcanic eruptions on Monsoon Asia's streamflow varies spatially, ranging from strong positive, mixed, to little. The mechanisms underlying this spatial variability are yet to be elucidated.

Our results are mostly consistent with Anchukaitis et al. (2010), who used Superposed Epoch Analysis to analyze PDSI anomalies in the eruption years. The key difference is in eastern China, where Anchukaitis et al. (2010) showed negative PDSI in year t, while we observed normal to positive streamflow anomalies in year t, and negative streamflow anomalies in year $t+1$ (Figure 4.7b). The discrepancies may be due to the different eruption data sets (Anchukaitis et al. (2010) demonstrated this with three sets of events) and the analytical methods. It could also be because they analyzed PDSI while we analyzed streamflow. That we observed negative streamflow anomalies in year $t+1$ instead of t could be due to the lagged response of streamflow in this region.

As a drought/pluvial indicator, streamflow may differ from PDSI in individual years for some regions, as discussed above, but on longer terms, our reconstructed streamflow agrees well with reconstructed PDSI. For example, our record fully captures the Angkor Droughts (1345–1374 and 1401–1425) (Buckley et al., 2010; Buckley et al., 2014) with prolonged low flow throughout Southeast Asia, and extended as far as India (Figure 4.6b, panels 2 and 3), in agreement with speleothem records from Dandak and Jhuma Caves (Sinha et al., 2007; Sinha et al., 2011). Heavy monsoon rain interrupted these megadroughts; such abrupt alterations to the flow regime proved difficult for the

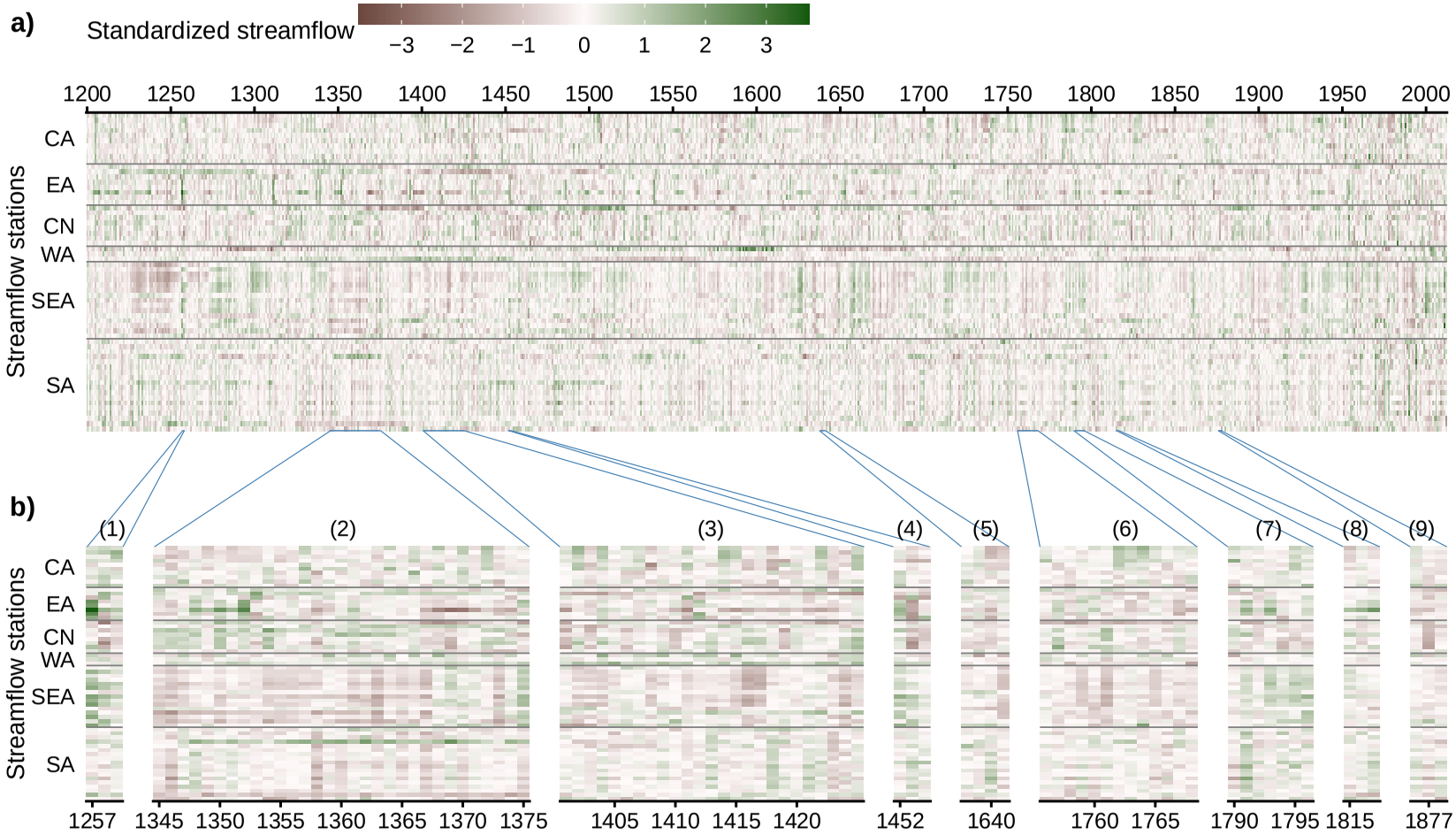

Figure 4.6. Spatiotemporal variability of streamflow in Monsoon Asia. a) Variations over time (x-axis) and space (y-axis) of the standardized streamflow index (i.e., z-score of streamflow, or z-score of log-transformed streamflow when log-transformation was used in the reconstruction). The stations are arranged approximately north to south (top down on y-axis) and divided into five regions as delineated in Figure 4.1: CA (Central Asia), EA (East Asia), WA (West Asia), CN (eastern China), SEA (Southeast Asia), and SA (South Asia). b) Historic events captured in the reconstruction: (1) Samalas eruption, (2) and (3) Angkor Droughts I and II, (4) Kuwae eruption, (5) Ming Dynasty Drought, (6) Strange Parallels Drought, (7) East India Drought, (8) Tambora eruption, and (9) Victorian Great Drought.

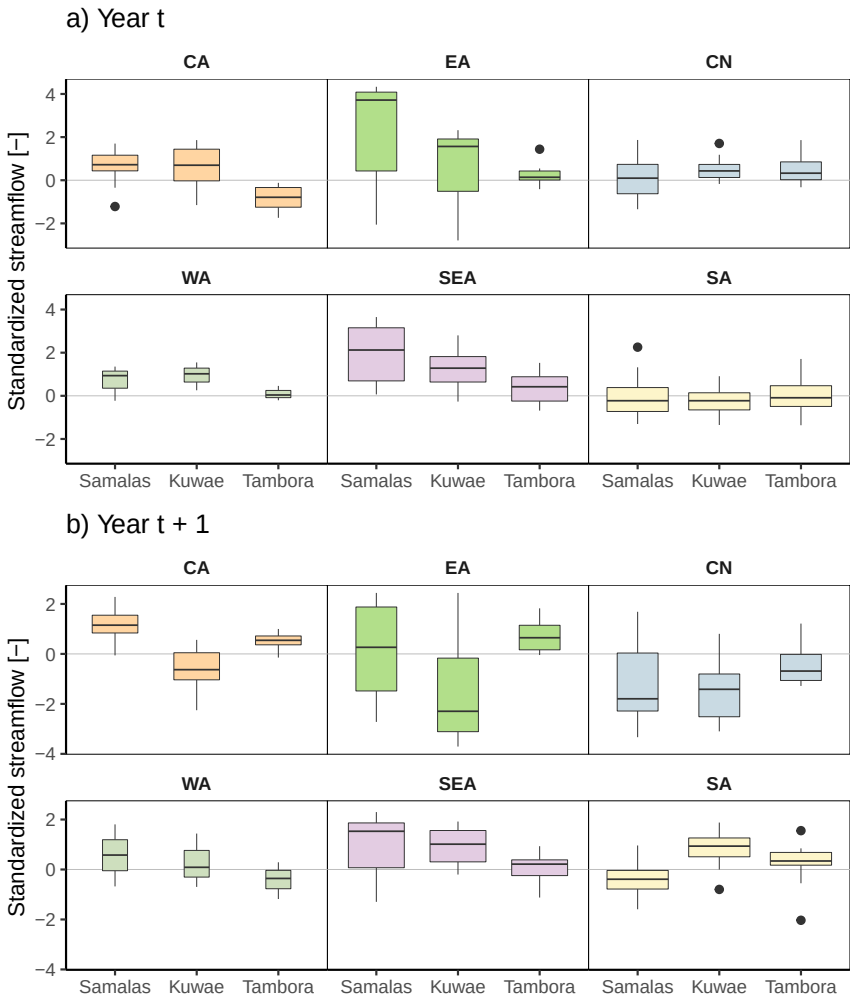

Figure 4.7. Distribution of standardized streamflow index in three volcanic eruptions. The widths of the box plots are proportional to their sample sizes.

ancient city of Angkor (Buckley et al., 2014). The city once thrived thanks to an extensive network of hydraulic infrastructure (Lieberman & Buckley, 2012). After the first Angkor Drought, the inflow/outflow functions of the *barays* (reservoirs) were altered in an attempt to preserve water. Heavy rains and flooding subsequently destroyed the reduced-capacity hydraulic infrastructure. This flood likely occurred in 1375 (Figure 4.6b, event 2). By the second Angkor Drought, the "hydraulic city" (Groslier, 1979; Lustig & Pottier, 2007) had insufficient water storage and could not recover.

Four more megadroughts that severely affected Asian societies (E. R. Cook et al., 2010a) are also captured in our reconstruction (Figure 4.6b, panels 5, 6, 7, and 9), along with other major droughts and pluvials. For example, Central Asia observed a six-decade drought between 1260–1320, and sustained pluvials during 1740–1769. Most notably, Southeast Asia suffered a drought between 1225–1255 that was comparable in length to

Angkor Drought I, but more severe in magnitude. Following this drought was a multi-decadal pluvial in 1271–1307. The drought is prominent in the speleothem record of J. K. Wang et al. (2019), and the pluvial can also be traced from there.

4.4.3 Links to Oceanic Drivers

To exemplify the spatial variation of how the oceans influence streamflow, we selected four river basins from west to east: Godavari, Chao Phraya, Mekong, and Yangtze, and selected one station from each basin. The selected stations are in the main stream and their reconstructions are statistically skillful.

We calculated the correlation between reconstructed annual streamflow at each station and the seasonal averages of global sea surface temperature (SST) for the period 1856–2012. The season definitions are: December to February (DJF), March to May (MAM), June to August (JJA), and September to November (SON). We also included JJA and SON of the prior year—JJA (-1) and SON (-1). Correlation patterns vary both seasonally and spatially, with differences among rivers and among oceans (Figure 4.8).

4.4.3.1 Pacific Ocean

Tropical Pacific SST correlates significantly with streamflow at all four basins, but the correlation patterns vary. For the Godavari, moderate positive correlations are seen from JJA (-1) to DJF, and strong negative correlations are seen from JJA to SON. For the Yangtze, the pattern is completely opposite: strong positive correlations from JJA (-1) to DJF, and moderate negative correlations in JJA and SON. The location of the strongest correlations suggests links to the El Niño-Southern Oscillation (ENSO, cf. McPhaden et al. (2006)). We find it interesting that ENSO seems to influence the Godavari and Yangtze in contrasting ways.

Unlike the Godavari and Yangtze, the Chao Phraya and Mekong's streamflow correlates significantly with SST over most of the Pacific Ocean, and the correlation persists across all seasons. The correlation pattern is negative in the tropical Pacific, and positive in the northern and southern Pacific. This pattern and its lack of seasonality suggest that, beside ENSO, there are influences from a driver at longer time scales, likely the Pacific Decadal Variability (PDV)—decadal variations of Pacific SST resulted from complex tropical-extratropical ocean-atmosphere interactions (Henley, 2017). The North Pacific component of PDV is known as the Pacific Decadal Oscillation (PDO) (Mantua & Hare, 2002), its southern counterpart the South Pacific Decadal Oscillation (Shakun & Shaman, 2009); basin-wide SST variation patterns have also been termed Interdecadal Pacific Oscillation (Folland et al., 1999). These modes are closely related (Henley, 2017). The PDV has been shown to influence hydroclimatic variability in Monsoon Asia, in conjunction with ENSO (Yu et al., 2018). Specifically for the Chao Phraya, PDV also modulates ENSO's influence on peak flow (C. Xu et al., 2019).

4.4.3.2 Indian Ocean

Correlation patterns are less prominent in the Indian Ocean compared to the Pacific. We observe basin-wide correlations in DJF for the Godavari and Yangtze; correlations bear

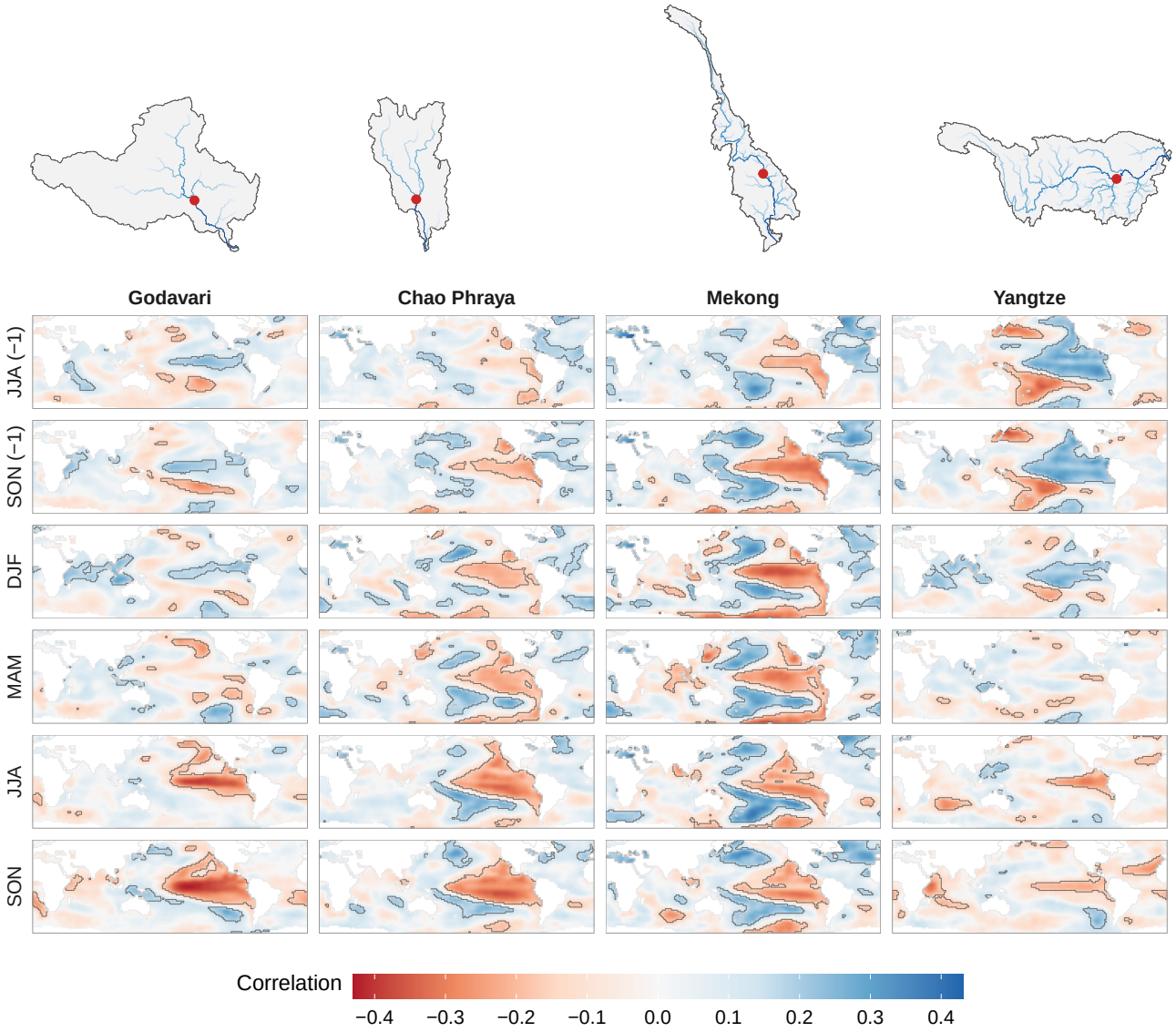

Figure 4.8. Correlation between reconstructed mean annual streamflow at four river basins (this work) and seasonal averages of global sea surface temperature (SST) from the NOAA_ERSST_v5 data set (Huang et al., 2017) for the period 1856–2012; significant correlations ($\alpha = 0.05$) enclosed in black boundaries. The locations of the stations are shown in the catchment maps; these are the same stations shown in Figure 4.3. Seasons are marked by the year in which they end. "(-1)" denotes previous year.

the same sign as that in the Pacific. This is consistent with the Indo-Pacific coupling: an ENSO event in the Pacific leads to SST anomalies of the same sign in the Indian Ocean (Saji et al., 1999). The Godavari and Yangtze also exhibit another correlation pattern in SON (with small areas of significance): correlations bear opposite signs between the tropical western Indian Ocean near the Horn of Africa and the southeastern Indian Ocean around Sumartra. This pattern and its timing suggest links to the Indian Ocean

Dipole (IOD) (Saji et al., 1999; Ummenhofer et al., 2017). The IOD accounts for about 12% of Indian Ocean SST variability, much less than the basin-wide coupling mode (30%) (Saji et al., 1999); this explains the weaker correlations of the IOD. Positive IOD events have also been linked to droughts in Southeast Asia, but this relationship is not robust (Ummenhofer et al., 2013). In our analysis, the link between IOD and Southeast Asian streamflow is not visible. Our interpretation is that ENSO and PDV are the main drivers here, and they dominate any links that the IOD might have.

4.4.3.3 Atlantic Ocean

The Chao Phraya and Mekong streamflow correlates positively with tropical and northern Atlantic SST. Significant and consistent correlations are observed throughout the seasons for the Mekong, but less consistent for the Chao Phraya. The link between tropical Atlantic SST and Southeast Asian hydroclimate was also found in a Laotian cave speleothem record (J. K. Wang et al., 2019). To explain this relationship, these authors examined SST, atmospheric pressure, and zonal moisture transport from climate model simulations, and proposed the following mechanism: increased tropical Atlantic SST leads to changes in zonal moisture transport, causing depression over tropical Indian Ocean, reducing rainout over the basin, leaving more moisture available to be transported to mainland Southeast Asia, ultimately strengthening Indian Monsoon rain over the region.

We repeated the correlation analysis above for other stations in the Godavari, Mekong, and Yangtze, where additional stations with statistically skillful results are available on the main stream. Results for those stations are consistent with what we report here (Appendix C.5).

4.4.3.4 Temporal variability of teleconnections

The correlation analysis of Figure 4.8 shows the spatial variation of the streamflow–SST teleconnection in Monsoon Asia. This analysis was done for the common period of SST and streamflow data (1856–2012). To explore whether and how the teleconnection patterns changed through time, we repeated the correlation analysis using a sliding 50-year window with 10-year increments. We show in Figure 4.9 three non-overlapping windows, and present all windows in Movie S1. Results show that all correlation patterns changed through time, echoing previous works that found non-stationarities in oceanic teleconnection (e.g., Krishna Kumar et al. (1999) and Singhrattna et al. (2005)). Correlations were much weaker in the period 1911–1960 compared to the preceding and subsequent five decades (Figure 4.9). Some patterns are more transient than others. The Yangtze's JJA-SON pattern of negative correlations with tropical Pacific was only strong in 1921–1980 (Movie S1). On the other hand, the Chao Phraya's SON positive correlations with tropical Pacific persisted throughout all periods. In 1901–1950, when ENSO teleconnection was the weakest for all rivers, tropical and northern Atlantic SST became the strongest teleconnection for the Chao Phraya and Mekong (Movie S1).

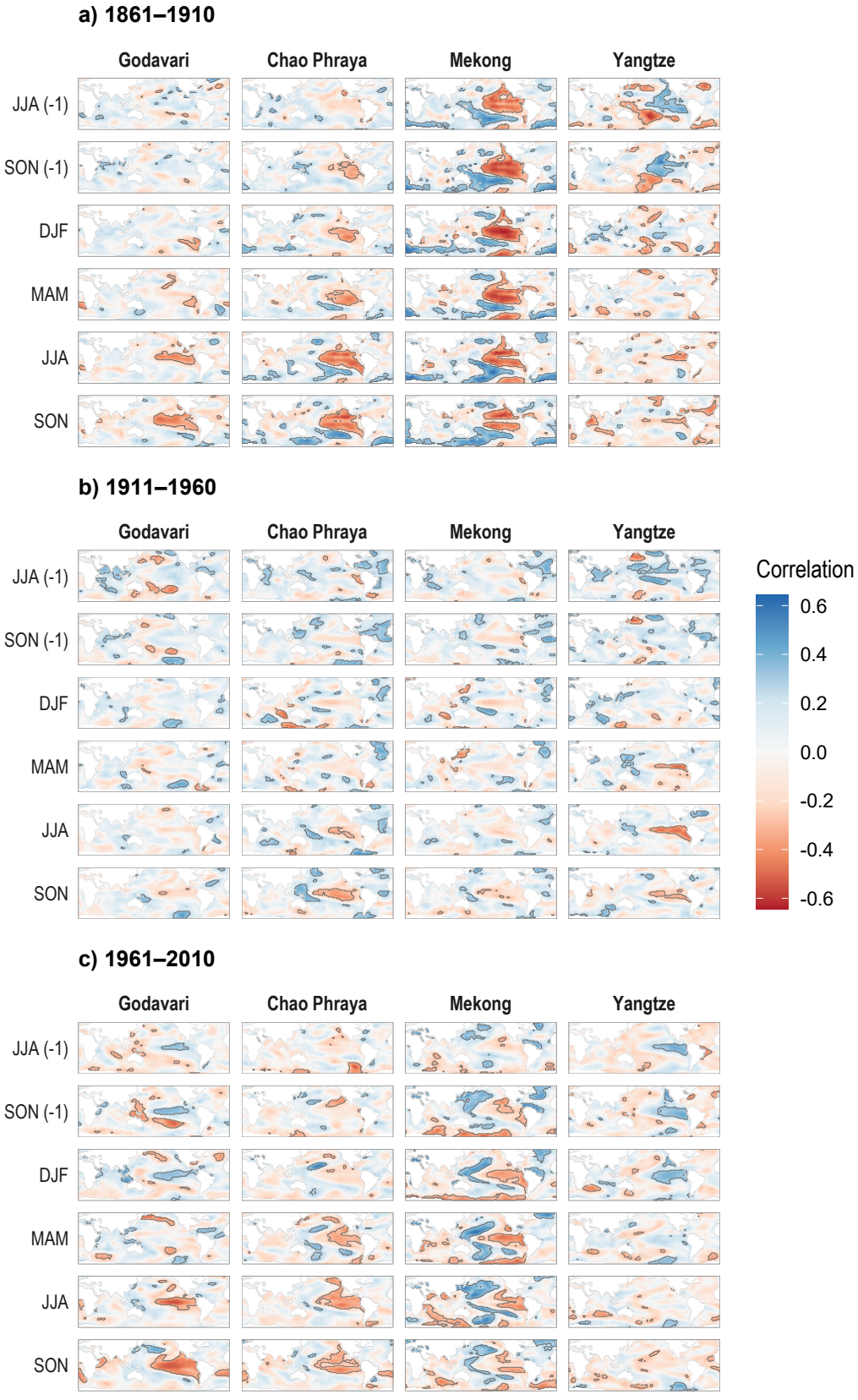

Figure 4.9. Temporal variability of the streamflow–sea surface temperature correlations. The analysis here is the same as that carried out in Figure 4.8, but split into three 50-year periods.

4.5 Conclusions

In this work, we produce the first large-scale and long-term record of streamflow variability for Monsoon Asia, covering 62 stations in 16 countries. In making this record, we also develop a novel automated, climate-informed, and dynamic streamflow reconstruction framework that leverages the computational advantages offered by our climate proxy—the Monsoon Asia Drought Atlas (MADA) version 2. Our framework achieves good skills for most of Monsoon Asia, and skill distribution is spatially homogeneous. Our results provide a regional, synthesized understanding of Monsoon Asia's streamflow variability over the past eight centuries, and reveal how the teleconnection between streamflow and its oceanic drivers varied over space and time.

From our reconstruction, streamflow in Monsoon Asia appears coherent: high and low flows often occur simultaneously at nearby stations and adjacent basins. This coherence is attributed to common oceanic drivers—the El Niño–Southern Oscillation (ENSO), the Pacific Decadal Variability (PDV), and sea surface temperature variations in the Indian and Atlantic Oceans. Coherence emerges even though we reconstructed each station individually, demonstrating the merits of Point-by-Point Regression. More importantly, this coherence implies that large-scale infrastructure transferring water, or other water-reliant commodities, across river basins could accidentally expose riparian people to unforeseen risks. For example, Thailand is increasingly purchasing Mekong-generated hydropower from Laos, and when that is insufficient, complements its energy needs with thermal power from plants that use water from the Chao Phraya for cooling. Thailand's energy system is more vulnerable when a prolonged drought occurs at both rivers (Chowdhury et al., 2020a)—our record shows such events have happened several times in the past.

We showed that the Pacific, Indian, and Atlantic Oceans influence streamflow variability, and that the strength and spatial footprint of these teleconnections varied over time. This result suggests that our understanding of how water-dependent infrastructure could perform may be narrow, especially in South and Southeast Asia, where we observe alternating periods of strong and weak teleconnections. A narrow characterization of climate-induced risks is likely to misguide climate change assessments, an important source of information for many major infrastructural decisions. Stakes are particularly high in Monsoon Asia, where river basins will experience further pressure in the coming decades (Satoh et al., 2017; Y. Wang et al., 2019). If we can develop methodologies for viewing future changes in streamflow in the context of past and present climate, we then have a pathway for making more informed and robust decisions. The reconstructions developed in our study offer a first step in this direction.

Chapter 5

A Probabilistic Assessment of the Chao Phraya Water System Using Reconstructed Monthly Inflows

5.1 Introduction

Rivers in their normal states support societies, but rivers in their extreme states—flood and drought—disrupt and may even destroy societies, e.g., the demise of Angkor due to alternating extreme droughts and floods (Buckley et al., 2010; Buckley et al., 2014). To cope with variations in streamflow, societies have built water management systems such as reservoirs, dams, and diversions. These systems are designed based on measured streamflow, which can capture only a limited range of streamflow variability (recall Figure 4.6). How will the system behave, and how will its performance be affected, when it experiences streamflow outside of what has been observed?

To answer this question, one can simulate the water system with reconstructed streamflow, which, thanks to the extended time span, may contain richer variability than do instrumental data alone. However, a typical annually resolved reconstruction does not provide the level of details needed to understand the system's behaviour. Furthermore, a historical streamflow time series, be it observed or reconstructed, is only one realization of the streamflow process, and provides only one snapshot of system performance. Therefore, water resources managers have used stochastic streamflow ensembles—suits of randomized time series that resembles the statistical properties of the observed time series—to study water systems probabilistically (Marco et al., 1993). Naturally, stochastic streamflow can also be generated for the reconstructed time series, creating paleoclimate-informed ensembles. One can generate stochastic ensembles of the annual streamflow reconstruction (Henley et al., 2011; Nguyen & Galelli, 2018) or stochastically disaggregate the annual reconstruction to higher frequencies (Prairie et al., 2008; Sauchyn & Ilich, 2017).

In this work, we test the idea of assessing a water system probabilistically with stochastic streamflow time series generated from monthly streamflow reconstruction. Our case study is the Chao Phraya water system (Thailand). We rely on the same proxies as used in Chapter 3, the Southeast Asian Dendrochronology Network, and reconstruct streamflow at the P.1 and N.1 stations on the Ping and Nan Rivers, both of which are major

tributaries of the Chao Phraya. The reconstructions are then used to generate stochastic streamflow ensembles using the Vector Autoregressive Moving Average model (Tsay, 2013). The ensembles are then used to simulate the water system. Finally, we assess the system performance with respect to three operating objectives: hydropower production, water supply, and flood control. This experimental study demonstrates the feasibility of doing probabilistic assessments of water system using monthly streamflow reconstruction. This experiment also reveals key challenges that limit the meaningful transfer from simulation results to actual operational improvements. We propose several research directions to overcome them.

5.2 The Chao Phraya Water System

The Chao Phraya River Basin has been introduced in Chapters 2 and 3. Here we focus on the Bhumibol and Sirikit Reservoirs, located on the Ping and Nan Rivers, respectively (Figure 5.1). These reservoirs are the two largest in Thailand; their characteristics are shown in Table 5.1. Both reservoirs are multi-purpose, serving irrigation, flood control, and hydropower production simultaneously.

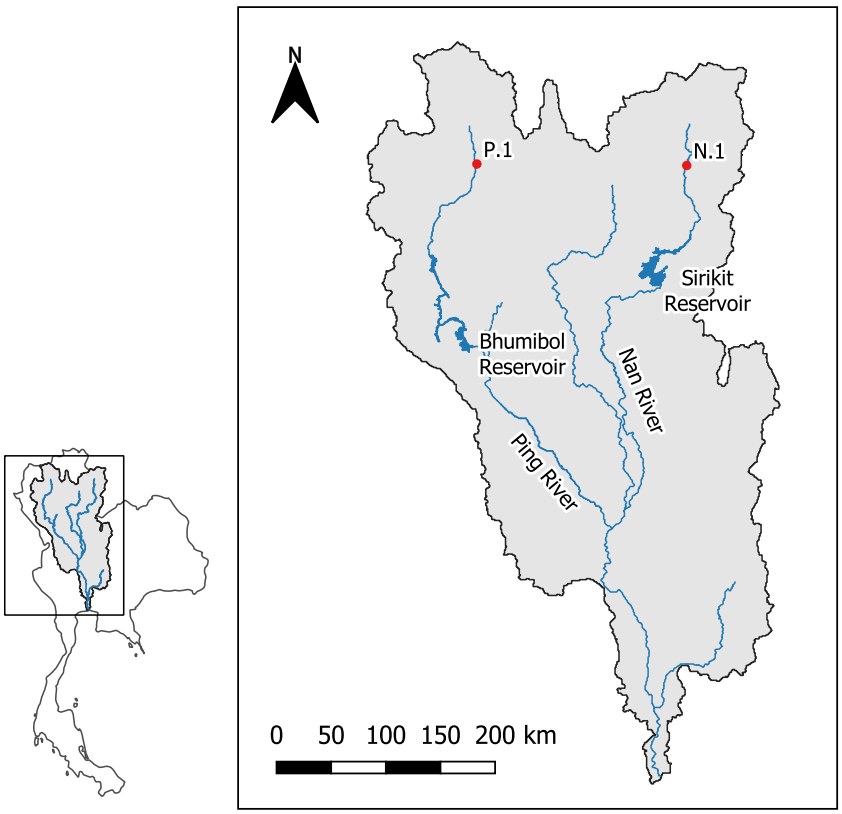

Figure 5.1. Map of the Chao Phraya River Basin, showing the two main tributaries (Ping and Nan), the streamflow stations P.1 and N.1, and the Bhumibol and Sirikit Reservoirs. The inset shows the location of the Chao Phraya River Basin in Thailand.

Table 5.1. Characteristics of the Bhumibol and Sirikit Reservoirs.

Attribute	Bhumibol	Sirikit
Opening year	1964	1974
Storage capacity (million m^3)	13,462	9,510
Active capacity (million m^3)	9,762	6,670
Dam height (m)	154	113.6
Installed capacity (MW)	779.2	500

Data on historical inflow, storage, and release are available at https://www.thaiwater.net/DATA/REPORT/php/rid_bigcm.php, but only from 2002. The short inflow time series are insufficient to use as reconstruction targets. Therefore, we reconstruct streamflow at the upstream stations P.1 and N.1 instead, and use a transfer function (Section 5.3.2) to convert upstream flows at P.1 and N.1 to inflow. Station P.1 has been used extensively in this thesis (Chapters 2, 3, and 4). Station N.1 were also included in the suit of reconstructions in Chapter 4. The streamflow measurements for both stations were obtained from the Thai Royal Irrigation Department (RID) at http://hydro-1.rid.go.th.

5.3 Methods

5.3.1 Monthly Reconstruction Model

Suppose we have S seasons, indexed as $s = 1, ..., S$. For example, in a monthly reconstruction, $S = 12$, and in a quarterly reconstruction, $S = 4$. In addition, we still have the annual reconstruction, indexed as q. The regression equations are then

$$\begin{aligned} \mathbf{y_s} &= \mathbf{X_s}\boldsymbol{\beta}_s + \boldsymbol{\varepsilon}_s \qquad \forall s = 1, ..., S \\ \mathbf{y_q} &= \mathbf{X_q}\boldsymbol{\beta}_q + \boldsymbol{\varepsilon}_q. \end{aligned}$$

Similarly to Section 3.3.1, let

$$\mathbf{y} = \begin{bmatrix} \mathbf{y_1} \\ \vdots \\ \mathbf{y_S} \\ \mathbf{y_q} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{X_1} & & & \\ & \ddots & & \\ & & \mathbf{X_S} & \\ & & & \mathbf{X_q} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_S \\ \boldsymbol{\beta}_q \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon} \\ \vdots \\ \boldsymbol{\varepsilon}_S \\ \boldsymbol{\varepsilon}_q \end{bmatrix}$$

and we can obtain back the canonical form of linear regression.

Finally, we need to expand the expression for δ:

$$\delta_t = \log\left(\sum_{s=1}^{S} \exp(\mathbf{x}_{\mathbf{s},t}\boldsymbol{\beta}_s)\right) - \mathbf{x}_{\mathbf{q},t}\boldsymbol{\beta}_q \quad \forall t = 1, ..., T$$

It is probable that some seasons do not need a log transformation while others do. In this case, we can define a transformation $t_s(.)$ for season s, and let $t_s^{-1}(.)$ be the inverse transformation. If there are no transformation for season s then both t_s and t_s^{-1} are the

identity transformation. We also define t_q and t_q^{-1} for the water year. The expression for δ is then

$$\delta_t = t_q \left(\sum_{s=1}^{S} t_s^{-1}(\mathbf{x_{s,t}} \boldsymbol{\beta_s}) \right) - \mathbf{x_{q,t}} \boldsymbol{\beta_q} \quad \forall t = 1, ..., T$$

The time series of the untransformed seasons can be orders of magnitude larger than those of the log-transformed seasons. In this case, all time series need to be standardized before the regression model can be trained. Standardization should also be applied even without log-transformation, when the wet seasons' flows are much larger than the dry seasons' flow.

In this study, we observed that with a monthly time step, the computational cost of solving the mass-balance-adjusted regression numerically is prohibitively high. Therefore, we opted to not log-transform any monthly time series so that the analytical solution can be used.

5.3.2 Reconstructing Reservoirs' Monthly Inflow

First, we reconstructed monthly streamflow at P.1 and N.1, using the model presented in Section 5.3.1. We then use the following transfer function to derive inflow to each reservoir from upstream flow (P.1 for Bhumibol and N.1 for Sirikit):

$$Q_{n,m}^k = a_m^k U_{n,m}^k + b_m^k \tag{5.1}$$

where Q is the inflow and U is the upstream flow; a and b are the regression coefficients; k indicates the reservoir (Bhumibol or Sirikit); n denotes the year, and m denotes the month. Equation 5.1 is calibrated using observed monthly upstream flow and inflow from January 2002 to October 2020.

We anticipate that reconstructed streamflow will have smaller variance than the observed streamflow, as seen with the seasonal reconstruction in Chapter 3. This "variance compression" problem is common in dendrohydrology, as trees do not capture saturation excess flow caused by "flashy" rainfall (Coulthard et al., 2016). Therefore, we produce a second set of reconstructions, where the reconstructed time series of each station is bias-corrected for the instrumental period using quantile mapping (Robeson et al., 2020; Gudmundsson et al., 2012). This is the same procedure that we used to naturalize streamflow at P.1 as presented in Chapter 3. While bias correction increases the variance of the reconstruction, it does not respect the annual mass balance. This is an important trade-off that requires further research.

5.3.3 Stochastic Inflow Regeneration

We generated stochastic streamflow time series for P.1 and N.1 using the Vector Autoregressive Moving Average (VARMA) model (Tsay, 2013, Chapter 3). The VARMA

stochastic process $\{\mathbf{y}_t\}$, $t = 1, ..., T$, is governed by Equation 5.2:

$$\mathbf{y}_t = \sum_{i=1}^{p} \boldsymbol{\phi}_p \mathbf{y}_{t-i} + \sum_{j=1}^{q} \boldsymbol{\theta}_j \boldsymbol{\varepsilon}_{t-j} + \boldsymbol{\varepsilon}_t \tag{5.2}$$

Here, $\mathbf{y}_t$ is a length-2 vector consisting of monthly streamflow at P.1 and N.1 at time t. The months are indexed as t from 1 to the total number of months in the time series; this is different from the index $m = 1, .., 12$ used in Equation 5.1. This is because VARMA does not model seasonality. As a result, we deseasonalized the time series by subtracting away the monthly means and dividing away the monthly standard deviations.

The VARMA parameters were fitted with a maximum likelihood procedure implemented in the R package MTS (Tsay & Wood, 2018). To determine the appropriate orders p and q, we tested all four combinations of $p = 1, 2$ and $q = 1, 2$. We found that $p = 1$ and $q = 1$ yielded the smallest Bayesian Information Criterion (BIC) score (Schwarz, 1978).

We then generated one hundred time series of deseasonalized streamflow at P.1 and N.1 using the fitted VARMA model. Each time series was then back-transformed using the appropriate monthly means and standard deviations, and finally converted to reservoir inflow using the model identified in Equation 5.1.

5.3.4 Modelling the Reservoir System

Upper and a lower rule curves are stipulated for each reservoir (Figure 5.2a). In reality, however, the reservoirs' storage levels are often outside the rule curves (Figure 5.2b). That is because water allocation in the Chao Phraya River Basin follows a complicated negotiation process (Takeda et al., 2016), which does not always respect the rule curves. It is difficult, if not impossible, to model these negotiations mathematically. As a result, we could not model exactly the historical operations of the reservoirs. Instead, we opt to simulate the operations of each reservoir according to a hypothetical rule curve, taken as the average of the upper and lower curves (Figure 5.2a).

The reservoir mass balance of reservoir k at time t is governed by Equation 5.3:

$$S^k_{t+1} = S^k_t + Q^k_t - R^k(S^k_t, Q^k_t) - s^k(S^i_t, Q^k_t) \tag{5.3}$$

where S is storage, Q is inflow, R is release, and s is spill. For simplicity, we shall drop the superscript k in the following discussion, where it is understood that each equation is parameterized differently for each reservoir.

Both R and s are functions of S and Q. Let L_t, M_t, and U_t be the lower, hypothetical (middle), and upper rule curves' target storage at time t. Then,

$$R(S_t, Q_t) = \begin{cases} 0 & \text{if } S_t + Q_t - R_{\min} < L_t \\ R_{\min} & \text{if } S_t + Q_t - R_{\min} \geq L_t \text{ and } S_t + Q_t < M_t \\ R_{\max} & \text{if } S_t + Q_t - M_t > R_{\max} \\ S_t + Q_t - M_t & \text{otherwise} \end{cases} \tag{5.4}$$

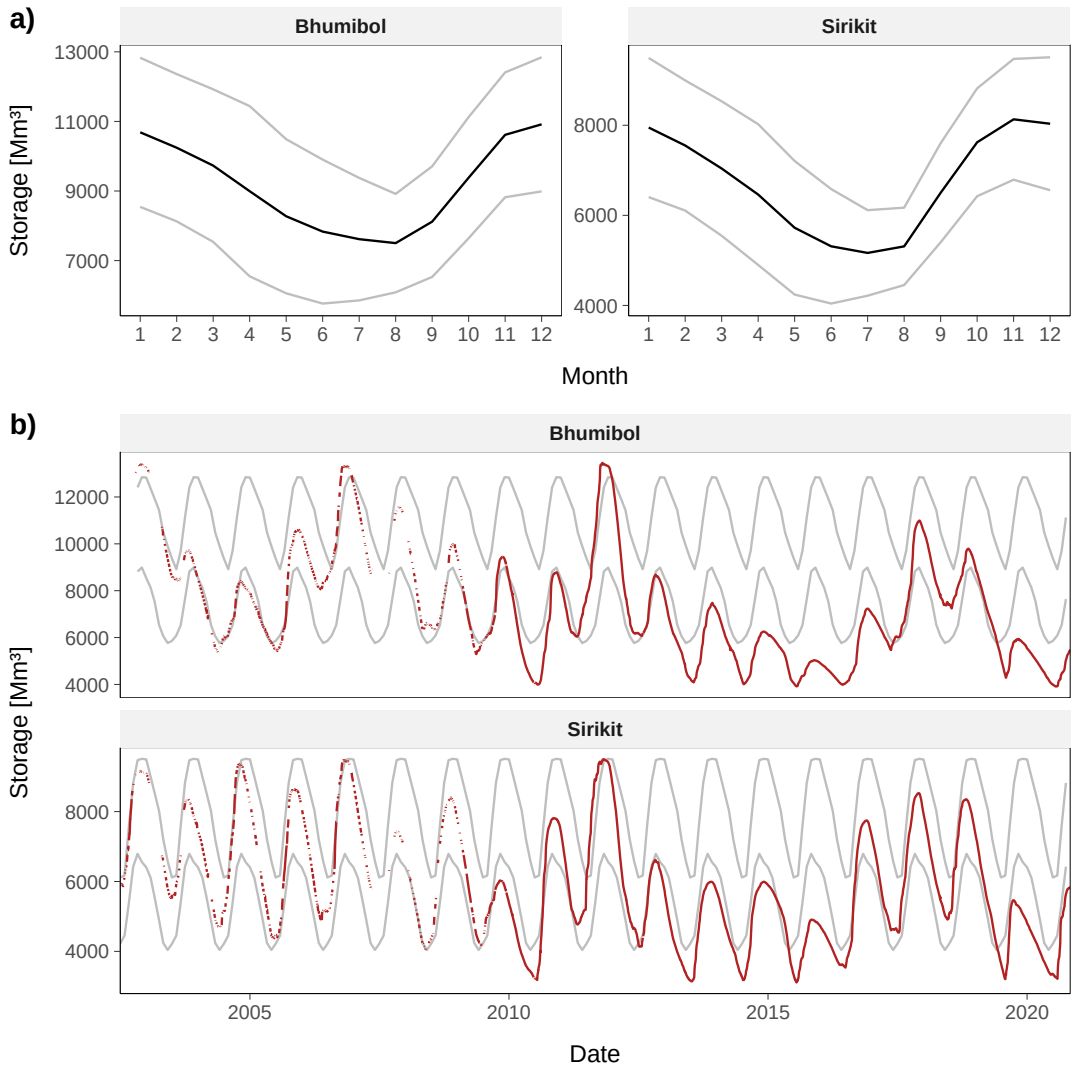

Figure 5.2. a) Upper and lower rule curves (gray) and hypothetical rule curve (black) of the Bhumibol and Sirikit Reservoirs. Bhumibol's rule curves were obtained from https://www.thaiwater.net/DATA/REPORT/php/rid_lgraph3.php?dam_id=1, and Sirikit's from https://www.thaiwater.net/DATA/REPORT/php/rid_lgraph3.php?dam_id=2. The hypothetical rule curve is the average of the upper and lower rule curves. b) Daily historical storage level from 2002 compared to the upper and lower curves.

and

$$s(S_t, Q_t) = \begin{cases} S_t + Q_t - R_{\max} & \text{if } R(S_t, Q_t) = R_{\max} \text{ and } S_t + Q_t - R_{\max} > U_t \\ 0 & \text{otherwise} \end{cases} \tag{5.5}$$

For each reservoir, $R_{\max}$ is determined from the turbine capacity, and $R_{\min}$ is taken as 1% of $R_{\max}$, which is close to the observed minimum release.

As we acknowledged earlier, the actual operations of the two reservoirs are complicated, resulting from negotiations that cannot be modelled mathematically. The operational characteristics that we adopted here is a much simplified representation. But, based on limited historical data available, we found that the our reservoir model is able to match the historical release reasonably well (Figure 5.3). Therefore, simulations using this model can be considered as a reasonable representation of the system performance.

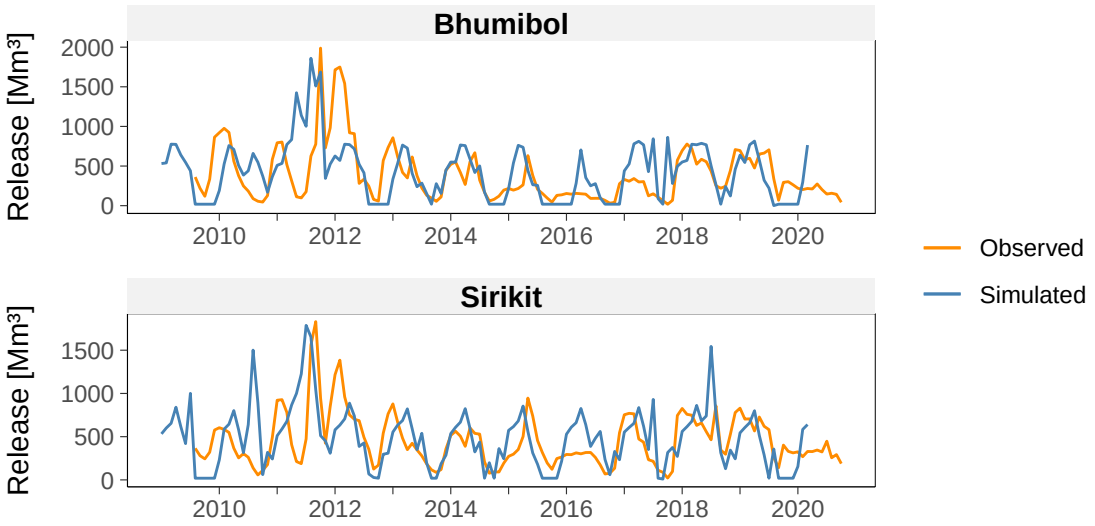

Figure 5.3. Comparison between simulated and historical release for each reservoir.

5.3.5 Experimental Setup

We simulate the Chao Phraya water system with three stochastic inflow ensembles. The first ensemble, called the "Modern" ensemble, is trained on the instrumental data; the second ensemble, "Paleo", is trained on the reconstructions; and the third ensemble, "BC Paleo", is trained on the bias-corrected reconstructions. Each ensemble has one hundred pairs of inflow time series for the two reservoirs. The time series are 254 years long, from 1750 to 2003, representing the reconstruction period. Each ensemble member produces one simulation run. Each simulation run yields the storage and release trajectories for the reservoirs corresponding to its inflow.

We evaluate the reservoirs' performance in each simulation with respect to the following metrics, which correspond to the three operating objectives of hydropower production, water supply, and flood control.

Mean annual hydropower production

$$H^k = \frac{1}{N}\sum_{n=1}^{N}\left(\sum_{m=1}^{12}\eta\rho g h(S^k_{n,m})R^k_{n,m}\right), \tag{5.6}$$

where $N = 254$ years is the total simulation length, $\eta = 0.9$ is the turbine efficiency, $\rho = 1000\,\mathrm{kg\,m^{-3}}$ is the density of water, $g = 9.81\,\mathrm{m\,s^{-2}}$ is the free fall acceleration, $h^k(S^k_{n,m})$ is the turbine head at storage $S^k_{n,m}$, and $R^k_{n,m}$ is the release.

Mean annual supply level

We assume that the release from each reservoir is used to first serve the irrigation area immediately downstream, and any remaining amount is then used to serve the main Chao Phraya irrigation area. Therefore, we define the portion of reservoir k's release that is available to the main basin, denoted $r^k_{n,m}$ as

$$r^k_{n,m} = \max\{R^k_{n,m} - d^k_m, 0\} \tag{5.7}$$

where d^k_m is the demand in month m from the irrigation area immediately downstream of reservoir k. The system's mean annual supply level is define as

$$W = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{12}\sum_{m=1}^{12}\min\left\{\frac{r^{Bhumibol}_{n,m} + r^{Sirikit}_{n,m}}{D_m}, 1\right\}\right) \tag{5.8}$$

where D_m is the demand in month m from the Chao Phraya's irrigation area. Recall that N is the simulation length (254 years). The monthly demands are derived from Divakar et al. (2011).

Number of spill events

As we did not reconstruct streamflow for the Wang River and Yon River tributaries, we cannot model the inundation in the Chao Phraya delta. However, we can use the number of spill events as a proxy for flood events. The number of spill event in each simulation is

$$N_s = \sum_{n=1}^{N}\sum_{m=1}^{12}\mathbb{1}_{\{s_{n,m}>0\}} \tag{5.9}$$

These metrics reflect the aggregated long term performance of the reservoirs. In future works they will be used to optimize the reservoir operations, and the optimized metrics can then be compared with the baseline results presented here.

5.4 Results

5.4.1 Monthly Reconstructions

The reconstructions capture normal and low flows well, but not peak flows (Figure 5.4). This is also reflected in the skill scores, where August and September have the lowest scores (Table 5.2). In general, the reconstructions can still be considered skillful, as all RE and CE values are positive.

As the reconstructions have smaller variances than the instrumental data (Figure D.2); this reduced variance is observed for both instrumental and paleo periods. Therefore, we standardized each station's reconstruction with its monthly means and standard

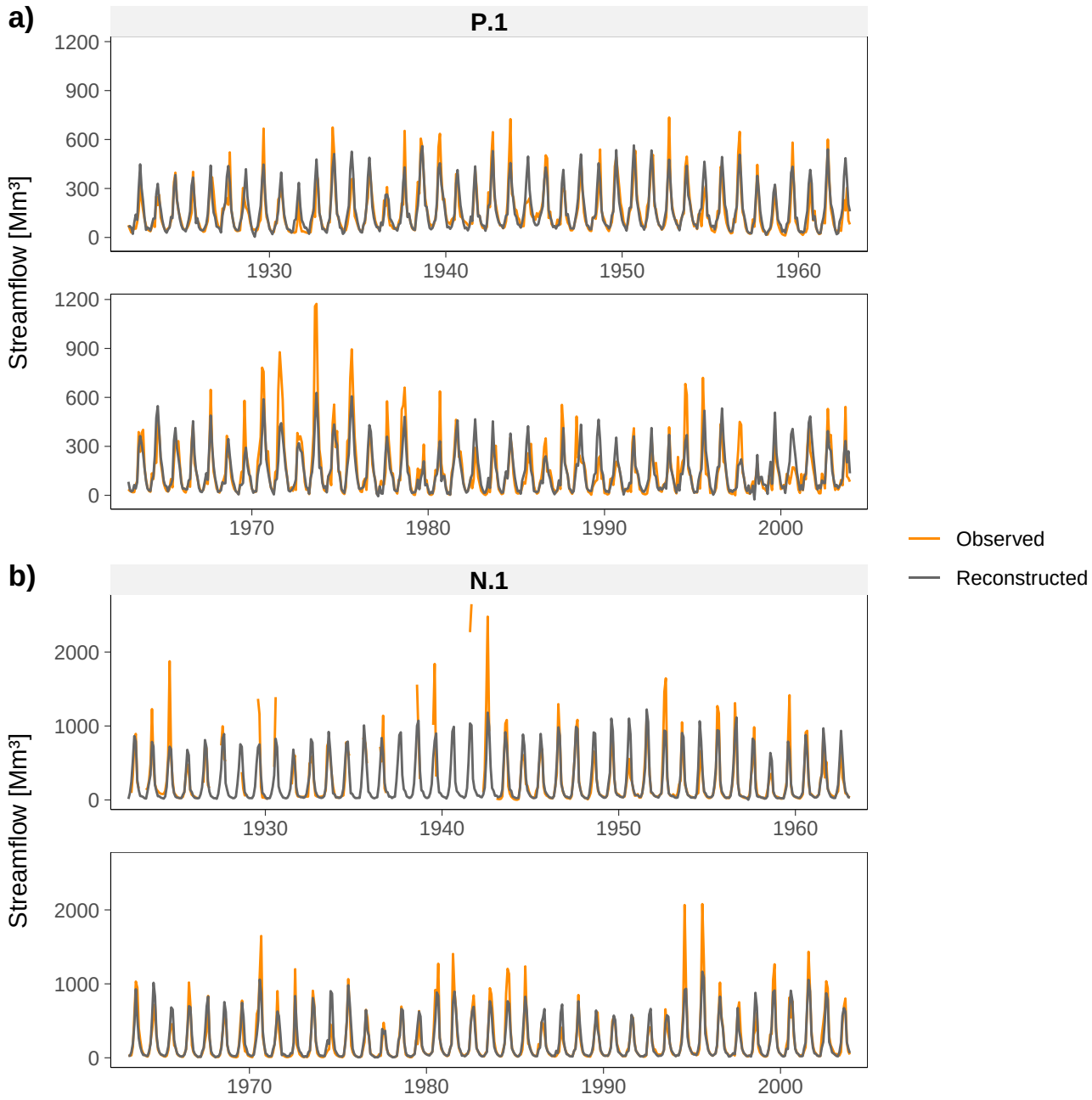

Figure 5.4. Reconstructions compared against observations.

deviations to examine the "internal" variability. The resulting z-scores reveal a rich history of streamflow variability that is quite different from that provided by the instrumental data (Figure 5.5). The distriubtions have different shapes: the instrumental data are more skewed. While not capturing the extreme high flows, the reconstructions capture more extreme low flows than do the instrumental data. This suggests that applying variance correction to the reconstructions will result in a much broader range of

Table 5.2. Reconstruction skill scores

	P.1			N.1		
Target	R^2	RE	CE	R^2	RE	CE
Jan	0.47	0.53	0.36	0.47	0.33	0.12
Feb	0.64	0.70	0.10	0.59	0.53	0.27
Mar	0.33	0.54	0.03	0.33	0.48	0.21
Apr	0.50	0.62	0.05	0.44	0.42	0.32
May	0.46	0.43	0.40	0.43	0.33	0.23
Jun	0.33	0.34	0.29	0.41	0.50	0.43
Jul	0.42	0.30	0.29	0.44	0.39	0.20
Aug	0.23	0.22	0.12	0.28	0.25	0.20
Sep	0.22	0.19	0.16	0.26	0.28	0.19
Oct	0.44	0.50	0.40	0.50	0.45	0.39
Nov	0.37	0.46	0.33	0.53	0.53	0.39
Dec	0.35	0.34	0.29	0.47	0.51	0.26
Annual	0.42	0.44	0.35	0.41	0.36	0.32

variability than what is observed in the instrumental data (recall Section 5.3.2).

A simple variance correction procedure commonly used in dendrohydrology is to scale the reconstruction by the ratio between the variance of the reconstruction's instrumental period and that of the instrumental data (e.g., D'Arrigo et al., 2011a). When applied to our reconstructions, this procedure produced too many negative flow values. Therefore, we resorted to using quantile mapping. While quantile mapping ensures that no negative flows are produced, it also makes the extreme low flow in the lower-left quadrant of Figure 5.5 become less extreme. The resulting distributions of the reconstructions match those of the instrumental data closely (Figure D.1).

5.4.2 Probabilistic Assessment

We first examine the storage behaviour of the reservoirs in the three stochastic ensembles. Only in the Paleo ensemble can the reservoirs maintain the target storage throughout the year (Figure 5.6). In the Modern and BC Paleo ensembles, both reservoirs can maintain the target storage only in May, June, and July. In August, the Bhumibol is often above the target, but the Sirikit is often below the target. In other months, the reservoirs are often below their targets. We acknowledge that the hypothetical targets that we set, which are the average of the upper and lower rule curves, are simplistic and may not have been the most appropriate for the reservoirs. For example, it appears that for Sirikit, the change from July's target to August's target is larger than the typical inflow, leading to Sirikit's storing less water than desired (and Bhumibol's situation is opposite). Nevertheless, these hypothetical targets serve to highlight the different behaviours of the reservoirs under different inflow scenarios. In future work we shall explore more methods to model the reservoir operations more accurately.

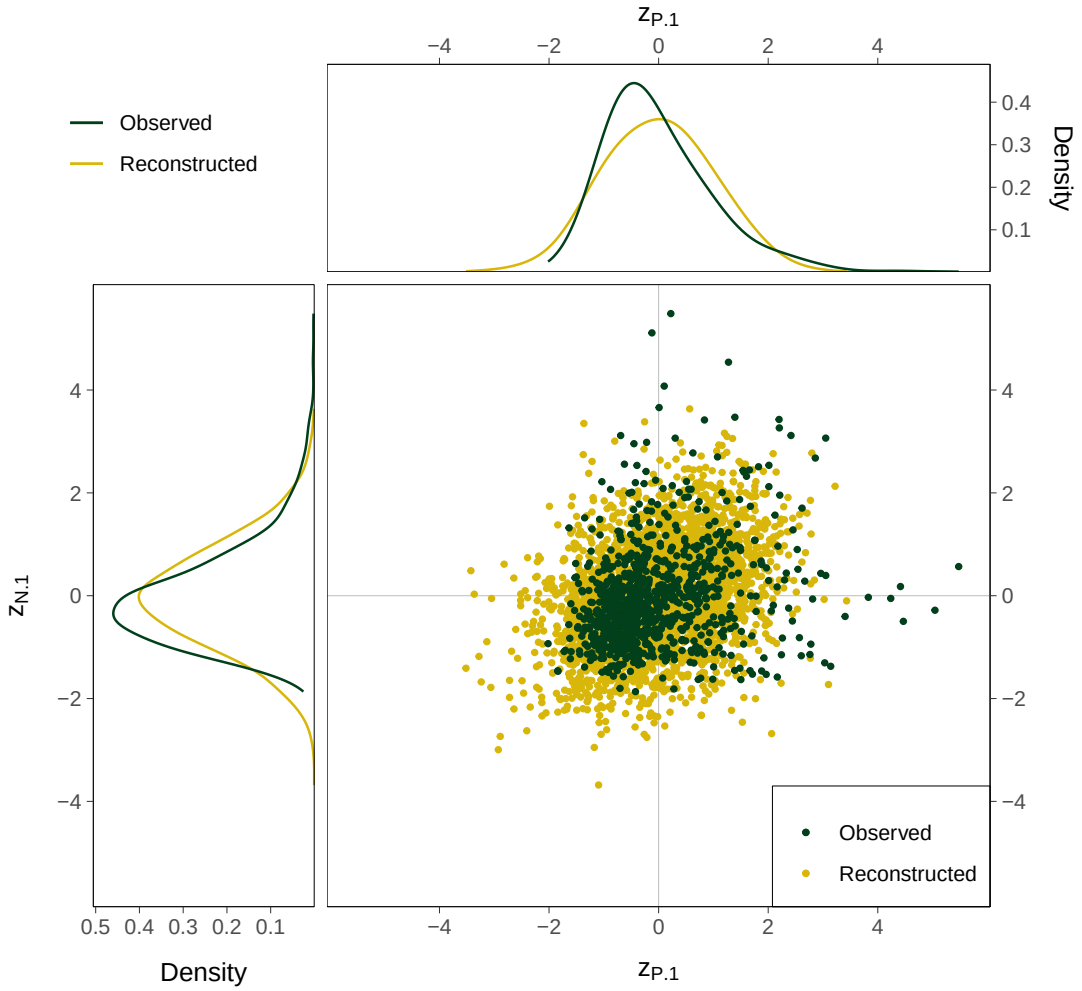

Figure 5.5. Marginal and joint distributions of observed and reconstructed standardized streamflow index. Streamflow values of each time series are converted to z-scores using the mean and standard deviation of that time series.

We now examine the distributions of the three performance objectives in the three stochastic ensembles. The results are shown in Figure 5.7.

Hydropower (Figure 5.7a)

Generally, the mean annual hydropower production, H, is higher in the Paleo ensemble than in the Modern ensemble (89/100 replicates for Bhumibol and 92/100 replicates for Sirikit). H also has a narrower spread in the Paleo ensemble. After the inflow reconstructions were bias-corrected with the BC Paleo ensemble, a much wider range of values were observed. Interestingly, in the BC Paleo ensemble, the Bhumibol observed more extremely low H, and the median H is about 60 GWh less than that in the Modern ensemble; on the other hand, the Sirikit observed more extremely high H, and the median H is about 27 GWh higher than that in the Modern ensemble.

These results can be explained by the storage behaviour that we observed in Figure 5.6.

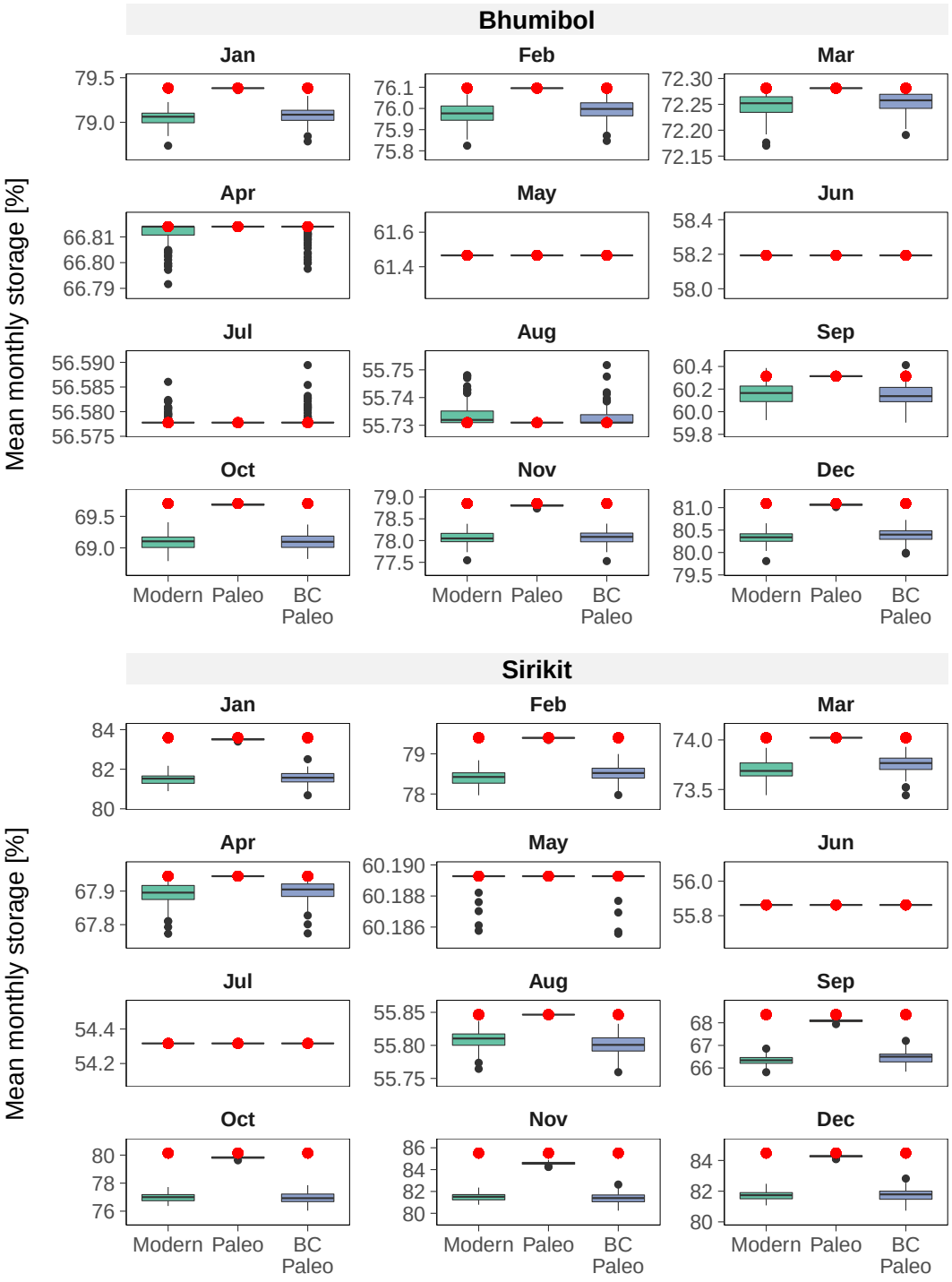

Figure 5.6. Mean monthly storage obtained from simulations with three stochastic ensembles: one trained on instrumental data (Modern), another on reconstructed inflow (Paleo), and the last on bias-corrected inflow reconstruction (BC Paleo). The mean monthly storage is calculated over 254 years for each ensemble member and for each month. The box plots show the distributions of the mean monthly storage over all members in each ensemble. The red dots denote the target storage.

In the Modern and BC Paleo ensembles, storage is often less than target, which leads to less release. Lower storage also means smaller hydraulic head. The compound effect of lower hydraulic head and less release causes hydropower production to drop.

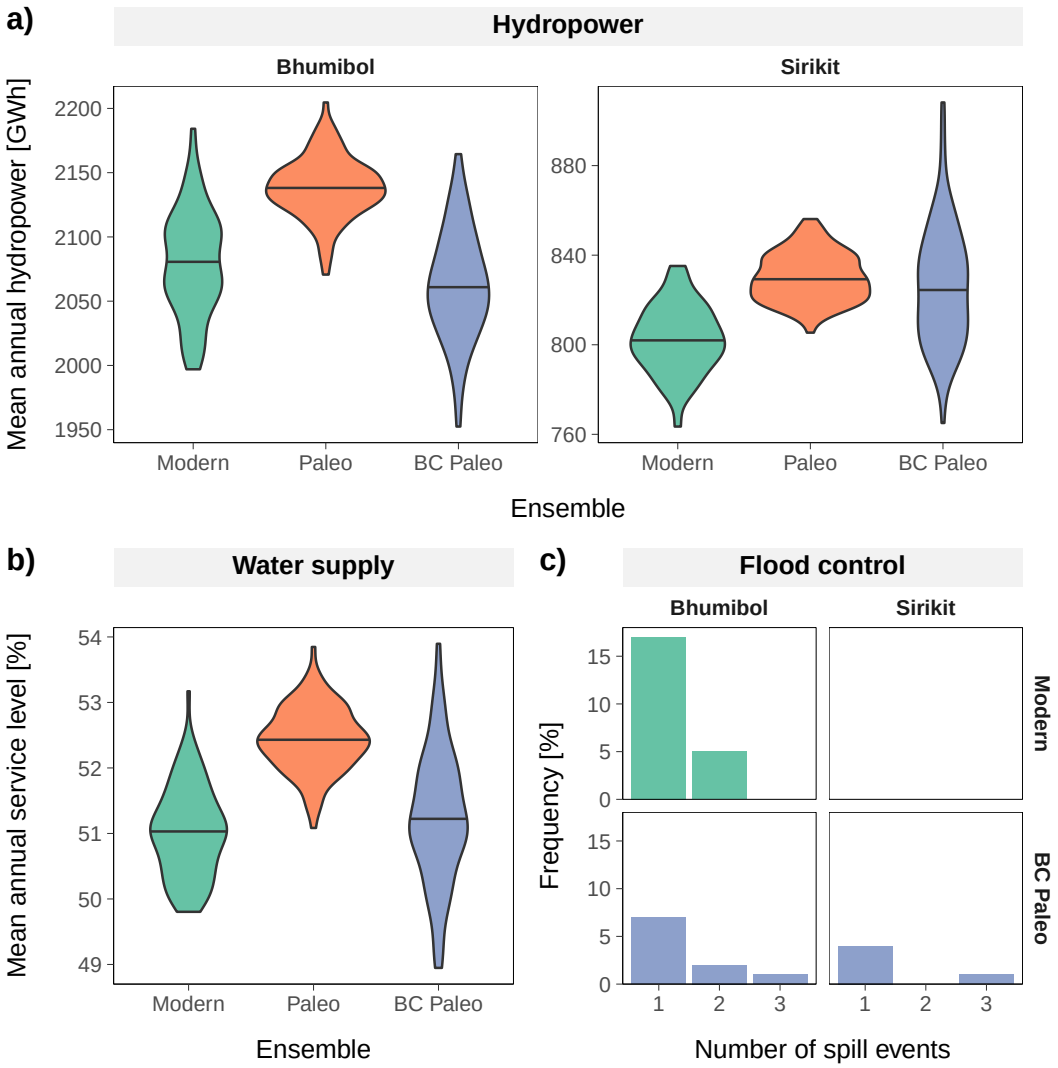

Figure 5.7. Distribution of three objectives (hydropower, power supply, flood controls) over three stochastic inflow ensembles. In panels (a) and (b), the black horizontal lines are the medians. In panel (c), the Paleo ensemble is not shown because it does not produce any spill events.

Water supply (Figure 5.7b)

The distributions of the mean annual service level, W, are similar to those of H. The system achieves about 1.5 percentage points higher W in the Paleo ensemble than in the Modern ensemble. The BC Paleo ensemble produces the largest spread, with the median W close to that in the Modern ensemble. These results can also be explained by the storage behaviour. Below-target storage levels, as observed in the Modern and BC Paleo ensembles, lead to supply cutoff. Above-target storage levels, however, does not

improve the service level beyond 100%, and therefore are not sufficient to compensate for the low supply months.

The Chao Phraya irrigation area is also served by the Wang and the Yon Rivers, in addition to the supply from the Bhumibol and Sirikit Reservoirs. The Wang and the Yon are not modelled in this study, but we aim to include them in future work.

Flood control (Figure 5.7c)

The Paleo ensemble does not produce any spill events. This is expected, as this ensemble is built directly from the reconstructions which do not capture extreme high flows. For the Bhumibol, 22 replicates in the Modern ensemble produce one or two spill events. Only 10 replicates in the BC Paleo ensemble do so, but one replicate produces three events. For the Sirikit, the Modern ensemble did not produce any spill events, but the BC Paleo ensemble has five replicates that have spill events, one of which has three events. All spill events occur in August, September, and October when the reservoirs receive the highest inflows.

5.5 Discussion and Conclusions

We presented monthly reconstructions for the P.1 and N.1 stations, which were then used to inform the stochastic inflow generator for the Bhumibol and Sirikit Reservoirs. With the skillful reconstructions, we tackled the main barrier to using streamflow reconstruction in water system operations: temporal resolution. Yet, our results point out another challenge—variance compression—which leads to an important trade-off. Not correcting for variance compression leads to a narrow range of inflow variability, from which system performance can be inflated. With variance correction, the range of variability is indeed increased, but the mass balance is worsened, which reduces the reliability of the results. The methods of variance correction also have their own limitations. Simple scaling leads to many negative flow values, while quantile mapping shifts the reconstruction's distribution and reduces the severity of extreme low flows.

While these challenges remain to be solved, our initial results affirm that when using paleo-reconstructed data, one gets a different range of system performance than when using instrumental data. These results show the promising role of streamflow reconstruction in improving water resources operations. But more work is needed to deliver that promise—the challenges identified here open up several research directions. Ideally, variance compression should be solved directly, by incorporating more proxies that can capture the extreme high flows. Alternatively, the linear dynamical system (LDS) model (Chapter 2) can be used. Both approaches require more computations, so future works need to address that. Meanwhile, a more appropriate variance correction method can be developed to capture the extreme high flow while maintaining the extreme low flow. Distributional discrepancies of instrumental and reconstructed data can also be homogenized using the Bayesian combination framework (Patskoski & Sankarasubramanian, 2015). These research directions can improve the accuracy and reliability of monthly streamflow reconstructions, such that in the future, they can be used, not just for probabilistic assessments, but for optimizing water resources systems under uncertainty.

Chapter 6

Synthesis

6.1 A New State of the Art and New Insights

The problem of managing the Colorado River's resources is a perfect motivating example for any class in water systems analysis. Its sheer size, socioeconomic context, and development history make it "one of the most controlled, controversial, and litigated rivers in the world" (Southern Nevada Water Authority, 2012). The Colorado River is also a classic case that motivates the study of dendrohydrology (Section 1.1). Even before the case of the Colorado River, dendrohydrology had been motivated by water resources management since as early as the 1930s (Section 3.1). Yet, the value of streamflow reconstructions in water management has been largely as a qualitative guidance, even in the discipline's birthplace—the United States. In Monsoon Asia, where streamflow reconstruction was developed much later and has been sparse in number, applications are even rarer, if at all.

The potential value of streamflow reconstructions has not been realized because of important gaps in the methodology and scope of applications. We identified four such gaps (Section 1.3): (1) most reconstructions rely on linear regression and do not capture catchment dynamics, (2) most reconstructions are annual, (3) few reconstructions have been developed in Monsoon Asia, all of which are single-sited—a synoptic understanding is lacking, and (4) reconstructions have not been used to study water system operations. This thesis contributes towards filling those gaps.

To account for catchment dynamics, I develop a novel reconstruction method based on linear dynamical systems (LDS), learned with a variant of the Expectation-Maximization (EM) algorithm (Chapter 2). The LDS model also serves as a stochastic generator of annual streamflow. The methodology has been implemented in the R package `ldsr`. Using the LDS model, we reconstruct streamflow for the Ping River (Thailand), unravelling four centuries of regime shifts in the catchment.

To achieve sub-annual reconstruction, I develop an optimal proxy selection procedure based on Genetic Algorithm, and a mass-balance-adjusted regression formulation to account for the annual mass balance (Chapter 3). This novel framework was used successfully to reconstruct seasonal streamflow for the Ping River, revealing interesting insights about the intra-annual streamflow variability while preserving the annual mass balance well (deviations were mostly within 10% of the mean annual flow). The framework has been implemented in the R package `mbr`.

To gain a synoptic understanding of streamflow variability across Monsoon Asia, I develop a large scale reconstruction, covering 62 stations in 16 countries (Chapter 4). We explore the use of a tree-ring-derived paleodrought atlas (the MADA), instead of the underlying tree ring network, as a paleoclimate proxy. The MADA offers several computational advantages to a large scale reconstruction across Monsoon Asia. The MADA grid points are selected for each station using a novel, climate-informed, and automated procedure to automate proxy selection. This procedure also involves a rigorous leave-k-out cross-validation scheme, and we advocate the use of this scheme instead of the traditional split-sample (two-fold) cross-validation. Our large scale reconstruction reveals that streamflow in Monsoon Asia is spatially coherent, owing to common oceanic drivers. The framework is being implemented in the R package `pprR`, which will soon be made available to the community.

To test, for the first time, the feasibility of using streamflow reconstruction in water system analysis (Chapter 5), I develop monthly reconstructions for the Ping and Nan Rivers and use them to build stochastic inflow ensembles to the Bhumibol and Sirikit Reservoirs. These ensembles are then used to simulate the performance of the Chao Phraya water system. While the monthly reconstructions are skillful, the probabilistic assessment of the water system proves difficult, mainly because of the variance compression phenomenon. Nevertheless, this experiment demonstrates that streamflow reconstructions, and their stochastic replicates, provide a different distribution of system performance than do instrumental data, reinforcing the need to incorporate streamflow reconstruction in future water system studies. When working on this chapter and related projects, I have also been putting together an R package called `shy` (**s**tochastic **hy**drology).

The research projects in this thesis have explored streamflow reconstructions from many angles, all aiming towards making streamflow reconstructions more valuable: streamflow reconstructions can now capture regime shifts, large scale variability, teleconnections, and intra-annual variability, and be used in operational studies of water system. The works also point out important open problems that call for further research.

6.2 The Next Steps

Variance compression remains the most important challenge. Particularly for the Ping River, this problem arises from the complex flood generation mechanism: floods often occur due to excess rainfall that fall on antecedent wet soil, either by typhoon rain after a heavy monsoon, or by monsoon rain after a typhoon (Lim & Boochabun, 2012; Stein et al., 2020). This mechanism cannot be captured by tree ring width. It was this mechanism that motivated us to explore the LDS model. Incorporating the LDS model into the monthly reconstruction framework in a computationally efficient way is an immediate direction.

To increase the reconstruction accuracy, more proxies could be incorporated. Stable oxygen isotope $\delta^{18}O$ has proved an important proxy not just for the wet season (Chapter 3). With recent technological advances (Gagen et al., 2011), it is becoming more affordable to extract $\delta^{18}O$ from tree rings. Therefore, we expect more chronologies to

be develop in Southeast Asia to contribute towards better streamflow reconstruction here. Blue intensity (Buckley et al., 2018) is also a promising proxy that can play a role in improving the reconstructions in the winter months.

Future dendrohydrological models may also build on a completely different approach: one could reconstruct time series of precipitation and temperature—instead of streamflow—and use them to force a process-based hydrological model, as we briefly discussed in Section 1.2. The approach requires a large amount of data for the model calibration, but it has two advantages. First, process-based models implicitly account for the physical consistency of the reconstructed discharge data. Second, they can represent alterations in the water cycle due to operations of hydraulic infrastructure or land use changes, enabling reconstructions for non-pristine catchments. This approach could be leveraged with state-of-the-art large-scale hydrological and water resource models enhanced with realistic water management representation.

A last important area, which is the main theme of my postdoctoral work, is the relation between dendrohydrology, water management, and the concept of nonstationarity. Dendrohydrological models build on the assumption that the relationship between proxies and discharge observed in instrumental records holds true over the entire length of the reconstruction. Such stationarity assumption has long been compromised not only by hydraulic infrastructure and land use changes, but also by anthropogenic climate change. In light of this expanded uncertainty, our community has progressively dropped the concept of the static design paradigm (Brown, 2010) and developed robust and dynamic planning techniques that adapt to nonstationary trends (Herman et al., 2020). So, shall we still use dendrohydrological data to plan and operate hydraulic infrastructure in a nonstationary environment? The answer is yes. Considering the case of anthropogenic climate change, we have at least two technical challenges in front of us. First, we shouldn't limit ourselves to compare the envelopes of reconstructed, observed, and projected streamflow variability. We need data analytics that quantify and reconcile the information contained in multiple power spectra and distributions. With such information, we could put future hydroclimatological risks into a better perspective, or develop a new generation of stochastic streamflow generators. Second, we should use the information provided by streamflow reconstructions to understand and, where possible, reduce the uncertainty in hydrologic projections informed by General Circulation Models (GCMs). A clear example is offered by the dynamics of the Asian Monsoon, which are well captured by climate proxies (Goodkin et al., 2019) but poorly represented by the latest generation of GCMs (Aadhar & Mishra, 2020). Streamflow reconstructions could therefore help us identify cases in which GCMs are not fully reliable, narrowing a major source of uncertainty for water resources management applications.

Epilogue

In my preliminary exam, after I presented Chapter 2 and my research plan, Prof. Peter Jackson, the committee chair, asked if I had ever seen tree rings in real life. Embarrassingly, I hadn't at the time (not in the scientifically meaningful sense; I had only seen tree rings on furniture). That question led me to attending the 29th Dendroecological Field Week in Croatia, where I cored my first trees (at the end of my third year!) The experience amplified my interests in tree ring science. One thing led to another, and now I have done less work in water management than I planned at the beginning of my PhD, but more reconstruction work than I had ever imagined.

I have been thinking a lot about the next step. On one hand, I will continue my dendrohydrological work, developing more methods to address the challenges that this thesis has identified, and gaining more insights about past hydroclimatic variability. On the other hand, I will experiment with more methods to incorporate paleoclimate reconstructions into practical applications. Overall, my research theme is to study hydroclimatic variability at all spatial and temporal scales, and translate that knowledge into water management. Improving water management has always been my goal, and dendrohydrology has become an important tool in my toolbox.

Appendix A

Supporting Materials for Chapter 2

A.1 Proof of Lemma 1

Proof.

Update A and B

From Equation 2.18:

$$\frac{\partial L}{\partial A} = \sum_{t=1}^{T-1} Q^{-1}(x_{t+1} - Ax_t - Bu_t)x_t'$$

Applying Leibniz's rule (swapping derivative and expectation) to Equation 2.19:

$$\frac{\partial \mathcal{Q}}{\partial A} = \mathbb{E}\left[\frac{\partial L}{\partial A}\right] = Q^{-1}\left(\sum_{t=1}^{T-1} P_{t+1,t} - A\sum_{t=1}^{T-1} P_t - B\sum_{t=1}^{T-1} u_t\hat{x}_t\right)$$

Set this partial derivative to zero:

$$A\sum_{t=1}^{T-1} P_t + B\sum_{t=1}^{T-1} u_t\hat{x}_t' = \sum_{t=1}^{T-1} P_{t+1,t} \tag{A.1}$$

Similarly,

$$\begin{aligned}\frac{\partial L}{\partial B} &= Q^{-1}\left(\sum_{t=1}^{T-1} \hat{x}_{t+1}u_t' - A\sum_{t=1}^{T-1} \hat{x}_t u_t' - B\sum_{t=1}^{T-1} u_t u_t'\right) \\ &\implies A\sum_{t=1}^{T-1} x_t u_t' + B\sum_{t=1}^{T-1} u_t u_t' = \sum_{t=1}^{T-1} \hat{x}_{t+1}u_t'\end{aligned} \tag{A.2}$$

Combining Equations A.1 and A.2:

$$\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} \sum_{t=1}^{T-1} P_t & \sum_{t=1}^{T-1} \hat{x}_t u_t' \\ \sum_{t=1}^{T-1} u_t \hat{x}_t' & \sum_{t=1}^{T-1} u_t u_t' \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{T-1} P_{t+1,t} & \sum_{t=1}^{T-1} \hat{x}_{t+1} u_t' \end{bmatrix}$$

$$\implies \begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{T-1} P_{t+1,t} & \sum_{t=1}^{T-1} \hat{x}_{t+1} u_t' \end{bmatrix} \begin{bmatrix} \sum_{t=1}^{T-1} P_t & \sum_{t=1}^{T-1} \hat{x}_t u_t' \\ \sum_{t=1}^{T-1} u_t \hat{x}_t' & \sum_{t=1}^{T-1} u_t u_t' \end{bmatrix}^{-1} \tag{A.3}$$

Update C and D

When there are no missing observations, all the y_t's are observed and treated as constants in the conditional expectation.

$$\frac{\partial \mathcal{Q}}{\partial C} = R^{-1} \left(\sum_{t=1}^{T} y_t \hat{x}_t' - C \sum_{t=1}^{T} P_t - D \sum_{t=1}^{T} u_t \hat{x}_t' \right)$$

$$\implies C \sum_{t=1}^{T} P_t + D \sum_{t=1}^{T} u_t \hat{x}_t' = \sum_{t=1}^{T} y_t \hat{x}_t' \tag{A.4}$$

$$\frac{\partial \mathcal{Q}}{\partial D} = R^{-1} \left(\sum_{t=1}^{T} y_t u_t' - C \sum_{t=1}^{T} \hat{x}_t u_t' - D \sum_{t=1}^{T} u_t u_t' \right)$$

$$\implies C \sum_{t=1}^{T} \hat{x}_t u_t' + D \sum_{t=1}^{T} u_t u_t' = \sum_{t=1}^{T} y_t u_t' \tag{A.5}$$

Combining Equations A.4 and A.5:

$$\begin{bmatrix} C & D \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{T} y_t \hat{x}_t' & \sum_{t=1}^{T} y_t u_t' \end{bmatrix} \begin{bmatrix} \sum_{t=1}^{T} P_t & \sum_{t=1}^{T} \hat{x}_t u_t' \\ \sum_{t=1}^{T} u_t \hat{x}_t' & \sum_{t=1}^{T} u_t u_t' \end{bmatrix}^{-1} \tag{A.6}$$

Update Q

$$\frac{\partial L}{\partial Q^{-1}} = \frac{T-1}{2}Q - \frac{1}{2}\sum_{t=1}^{T-1}(x_{t+1} - Ax_t - Bu_t)(x_{t+1} - Ax_t - Bu_t)'$$
$$\frac{\partial \mathcal{Q}}{\partial Q^{-1}} = \frac{T-1}{2}Q - \frac{1}{2}\left(\left(\sum_{t=1}^{T-1} P_{t+1} - AP_{t,t+1} - Bu_t\hat{x}'_{t+1}\right) + F_1 + F_2\right) \tag{A.7}$$

where

$$F_1 = \sum_{t=1}^{T-1}(AP_t + Bu_t\hat{x}_t - P_{t+1,t})A'$$
$$F_2 = \sum_{t=1}^{T-1}(Ax_tu'_t + Bu_tu'_t - \hat{x}_{t+1}u'_t)B'$$

Note that the values for A and B here are the updated values, i.e., those obtained in Equation A.3, which satisfy Equations A.1 and A.2. Therefore, $F_1 = F_2 = 0$, and Equation A.7 leads to

$$Q = \frac{1}{T-1}\sum_{t=1}^{T-1}(P_{t+1} - AP_{t,t+1} - Bu_t\hat{x}'_{t+1}) \tag{A.8}$$

Update R

By doing similar algebraic manipulations, we arrive at

$$\frac{\partial \mathcal{Q}}{\partial R^{-1}} = \frac{T}{2}R - \frac{1}{2}\left(\left(\sum_{t=1}^{T}(y_t - C\hat{x}_t - Du_t)y'_t\right) + G_1 + G_2\right) \tag{A.9}$$

where

$$G_1 = \sum_{t=1}^{T}(CP_t + Du_t\hat{x}'_t - y_t\hat{x}_t)C' \tag{A.10}$$
$$G_2 = \sum_{t=1}^{T}(C\hat{x}_tu'_t + Du_tu'_t - y_tu'_t)D' \tag{A.11}$$

Again, the values for C and D are the updated values, i.e., those obtained in Equation A.6, which satisfy Equations A.4 and A.5. Therefore, $G_1 = G_2 = 0$ and

$$R = \frac{1}{T}\sum_{t=1}^{T}(y_t - C\hat{x}_t - Du_t)y'_t \tag{A.12}$$

Update the initial state distribution (μ_1, V_1)

It is easy to show that $\mu_1 = \hat{x}_1$ and $V_1 = \hat{V}_1$.

■

A.2 Proof of Lemma 2

Proof. Let $\mathcal{O}$ be the set of time steps with observations and $\mathcal{M}$ be the set of time steps without observations. Then, define

$$Y_{\mathcal{O}} \coloneqq \{y_t : t \in \mathcal{O}\}$$
$$Y_{\mathcal{M}} \coloneqq \{y_t : t \in \mathcal{M}\}$$

There are two ways to specify the hidden data. The first and obvious way (approach 1) is to let the hidden data be $(X, Y_{\mathcal{M}})$ so that the complete data will be $(X, Y_{\mathcal{M}}, Y_{\mathcal{O}})$. However, the system equations implies that $Y_{\mathcal{M}}$ is independent of $Y_{\mathcal{O}}$ given X; therefore, we can specify the hidden data to be X and the complete data to be $(X, Y_{\mathcal{O}})$ (approach 2). Here we derive the update equations for approach 1 and show that it is indeed equivalent to approach 2.

First, we note that the conditional distribution for the hidden variable is now $p(X, Y_{\mathcal{M}}|Y_{\mathcal{O}}, \theta_k)$. Therefore, when we take expectation to obtain

$$\mathcal{Q} = \mathbb{E}_{X,Y_{\mathcal{M}}|Y_{\mathcal{O}},\theta_k}[\log p(X, Y_{\mathcal{M}}, Y_{\mathcal{O}}|\theta)],$$

we have to note that $Y_{\mathcal{M}}$ are random variables while $Y_{\mathcal{O}}$ are constants (in the full data case, all the y's are constants). As a result, the following quantities will arise for $t \in \mathcal{M}$

$$\mathbb{E}[y_t x_t'] = \mathbb{E}[(Cx_t + Du_t + v_t)x_t'] = CP_t + Du_t\hat{x}_t' \tag{A.13}$$
$$\begin{aligned}\mathbb{E}[y_t y_t'] &= \mathbb{E}[(Cx_t + Du_t + v_t)(Cx_t + Du_t + v_t)'] \\ &= (C\hat{x}_t + Du_t)(C\hat{x}_t + Du_t)' + C\hat{V}_t C' + R\end{aligned} \tag{A.14}$$

In both Equations A.13 and A.14, we used the fact that $x_t \perp v_t$. Furthermore in Equation A.14 we used

$$\mathbb{E}[w_t w_t'] = \mathbb{E}[w_t]\mathbb{E}[w_t'] + \mathbb{V}[w_t] = R$$

Second, since A, B, Q, μ_1 and V_1 are independent of Y, we only need to consider C, D and R.

Update C and D

Applying the above expectations, Equation A.4 becomes

$$\begin{aligned}&\sum_{t=1}^{T} P_t + D\sum_{t=1}^{T} u_t\hat{x}_t' = \sum_{t\in\mathcal{O}} y_t\hat{x}_t' + \sum_{t\in\mathcal{M}} CP_t + Du_t\hat{x}_t' \\ \implies\ & C\sum_{t\in\mathcal{O}} P_t + D\sum_{t\in\mathcal{O}} u_t\hat{x}_t' = \sum_{t\in\mathcal{O}} y_t\hat{x}_t'\end{aligned} \tag{A.15}$$

and Equation A.5 becomes

$$\begin{aligned} & C\sum_{t=1}^{T}\hat{x}_t u_t' + D\sum_{t=1}^{T}u_t u_t' = \sum_{t\in\mathcal{O}}^{T} y_t u_t' + \sum_{t\in\mathcal{M}}^{T}(C\hat{x}_t + Du_t)u_t' \\ \implies & C\sum_{t\in\mathcal{O}}\hat{x}_t u_t' + D\sum_{t\in\mathcal{O}}u_t u_t' = \sum_{t\in\mathcal{O}} y_t u_t' \end{aligned} \tag{A.16}$$

As a result, the new update equation for C and D is

$$\begin{bmatrix} C & D \end{bmatrix} = \begin{bmatrix} \sum_{t\in\mathcal{O}} y_t\hat{x}_t' & \sum_{t\in\mathcal{O}} y_t u_t' \end{bmatrix} \begin{bmatrix} \sum_{t\in\mathcal{O}} P_t & \sum_{t\in\mathcal{O}} \hat{x}_t u_t' \\ \sum_{t\in\mathcal{O}} u_t\hat{x}_t' & \sum_{t\in\mathcal{O}} u_t u_t' \end{bmatrix}^{-1} \tag{A.17}$$

which is identical to Equation A.6 except that the sum is over the observed time steps instead of over the whole time series length.

Update R

Similarly, substituting Equations A.13 and A.14 into the expectation of Equation A.9, we also obtain a new update equation for R as follows

$$R = \frac{1}{|\mathcal{O}|}\sum_{t\in\mathcal{O}}(y_t - C\hat{x}_t - Du_t)y_t' \tag{A.18}$$

where $|\mathcal{O}|$ denotes the number of elements in the set $\mathcal{O}$.

Thus we have shown that $Y_{\mathcal{M}}$ does not affect the M-step and it is sufficient to consider X as the hidden data. ■

Appendix B

Supporting Materials for Chapter 3

B.1 Imputing Tree Ring Data

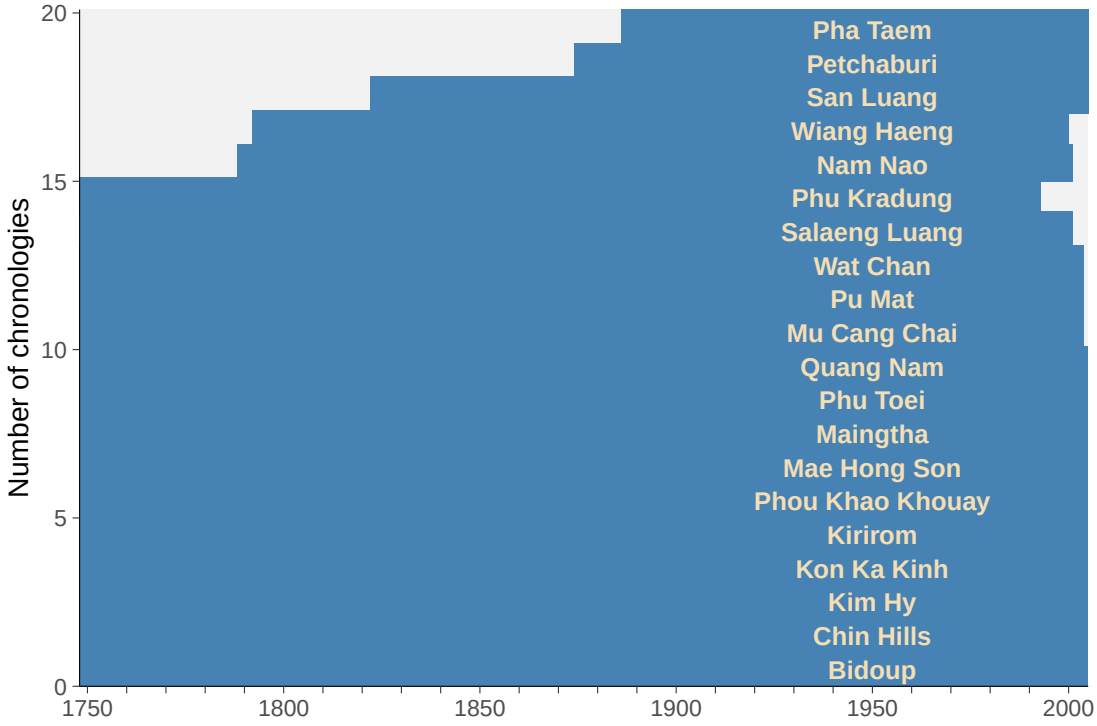

Figure B.1. Time spans of the tree ring width chronologies (blue areas). The gray areas are to be imputed; they accounts for about 8% of the total number of possible data points.

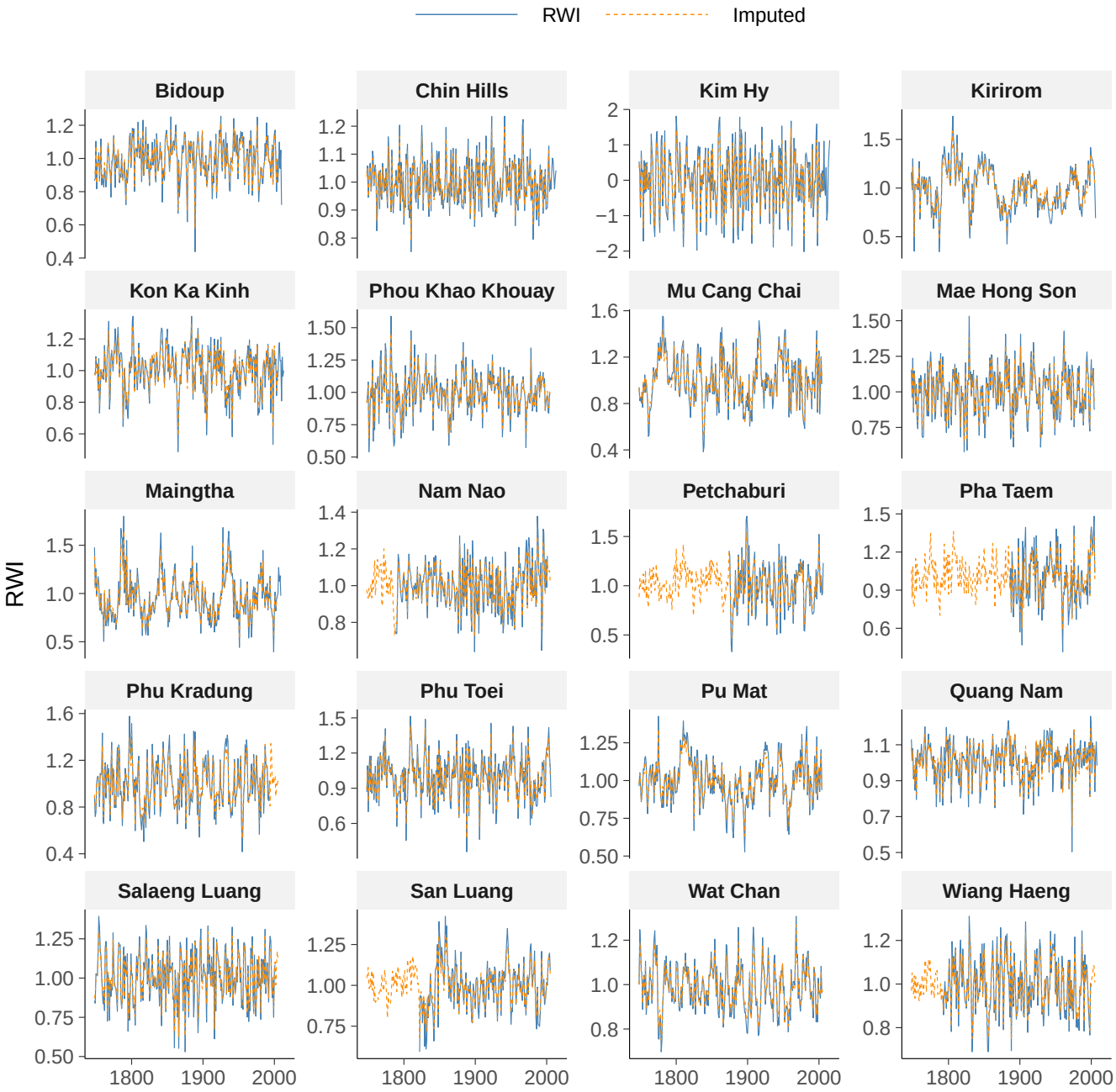

Figure B.2. Results of the imputation process. The original ring width indices (RWI) and the imputed values match very closely in the common periods. The imputations are applied to extend the chronologies backward and forward as necessary. Finally, we combine the imputed values in the missing years with the original ring width indices.

B.2 Streamflow Data Preprocessing

B.2.1 Gap Filling for Streamflow Data

Daily streamflow at P.1 is almost complete: from April 1921 to March 2017, there were only 87 missing days (0.25% missing rate). We simply filled each missing value with the average flow of the same date. The minimal fraction of missing data does not affect the seasonal flow calculation.

B.2.2 Season Delineation

In Thailand, the water year is considered to be the same as the Thai calendar year, from April to March. This definition does not match the hydrograph at P.1 very well (Figure B.3). Therefore, we delineate the wet season and dry season empirically using the procedure of B. I. Cook and Buckley (2009). For each year, we first calculated the daily cumulative streamflow (Figure B.4a), then we used the two-phase linear regression method (Lund & Reeves, 2002) to determine the change points in the slope of the cumulative streamflow time series. These change points are the dates of onset and withdrawal of the wet season in each year. Based on the trajectories and distributions of the change points (Figure B.4b and B.4d), we delineate the wet season as July–October, and the dry season as November–June.

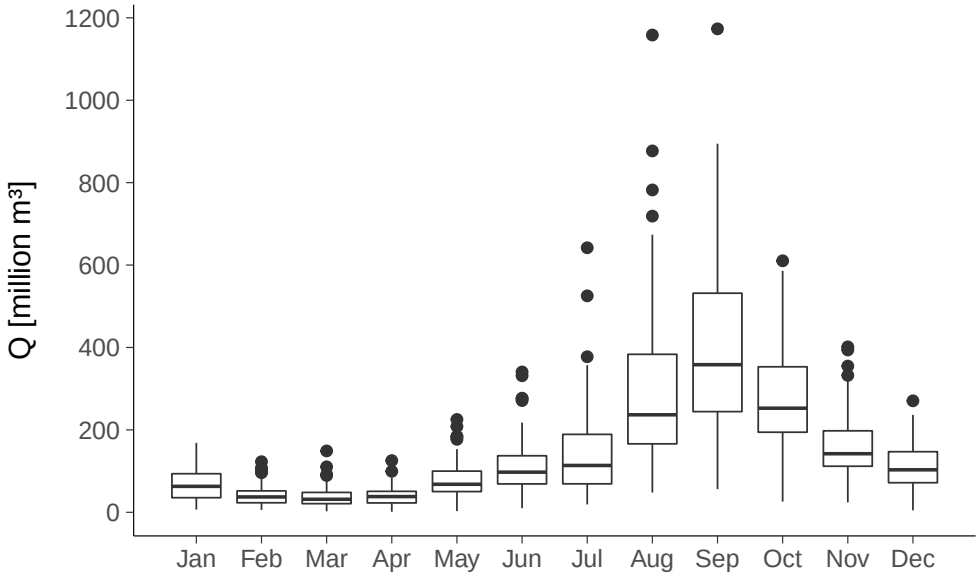

Figure B.3. Distribution of monthly streamflow at P.1 from instrumental data.

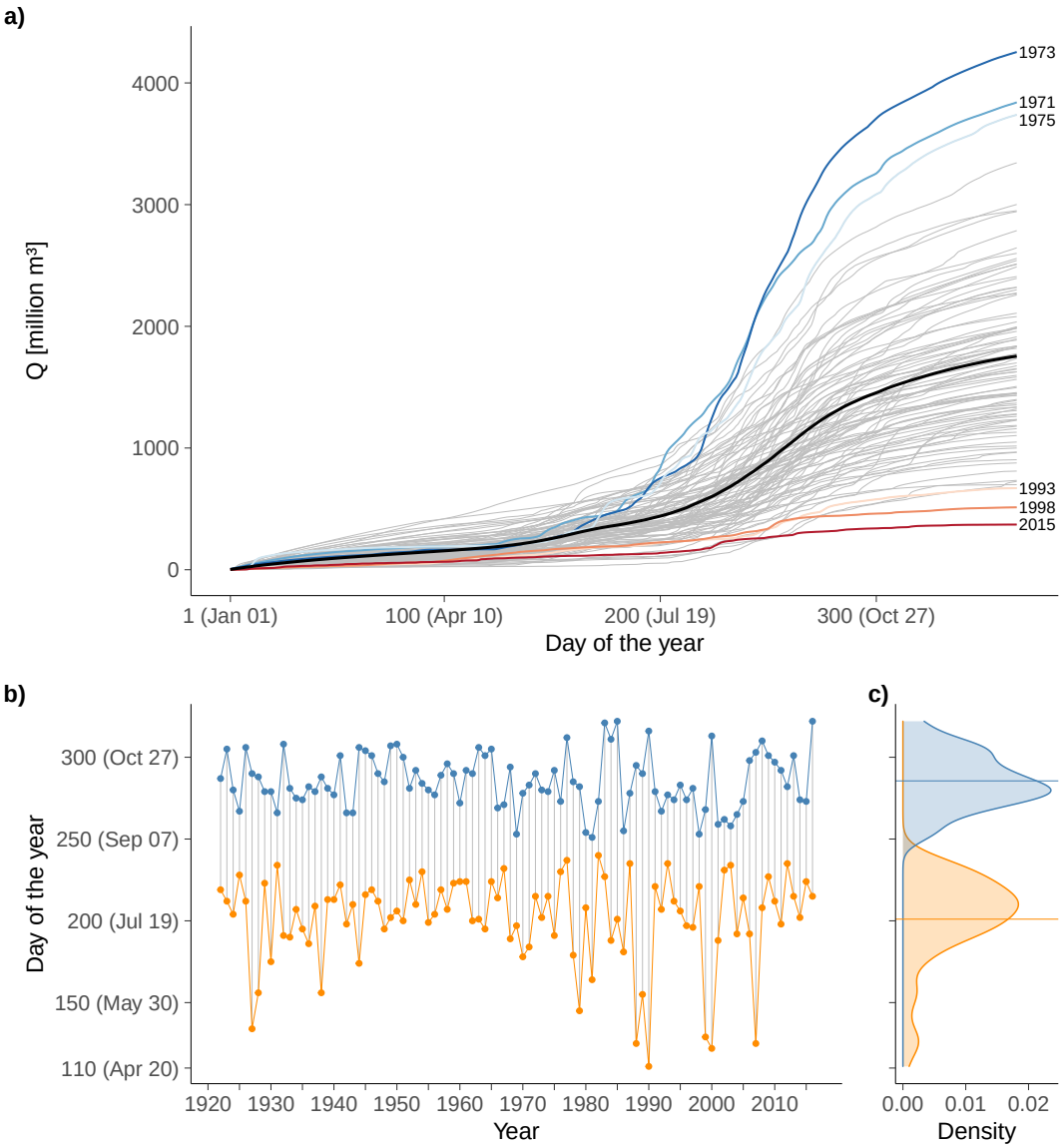

Figure B.4. Season delineation using the change point method. a) Cumulative daily discharge for each year. The three wettest years (blue) are La Niña years, and the three driest years (red) are El Niño years. The thick black line is the average. b) Dates of onset (orange) and withdrawal (blues) of the wet season in each year. c) Distributions of the onset and withdrawal dates.

B.2.3 Streamflow Naturalization

The ideal way to naturalize a streamflow time series is to build on observed data of the upstream anthropogenic interventions. For example, the case of a river segment affected by the operations of an upstream reservoir could be tackled by using data on inflow, operating rules, and release (Sauchyn & Ilich, 2017). Unfortunately, this information is not available for the Mae Ngat Dam, so we proceeded by relying

on discharge data simulated at station P.1 by VIC, a spatially-distributed, hydrologic-hydraulic model.

VIC is a popular large-scale model that simulates the water and energy balances at the land surface. Specifically, VIC organizes the spatial domain into a number of computational cells, where baseflow, runoff, infiltration, and evapotranspiration are estimated as a function of various hydro-meteorological forcings (Liang et al., 1994). The simulated runoff is then routed through the river network by a post-processor, or flow routing scheme (Lohmann et al., 1996; Lohmann et al., 1998). In our study, we use a model implementation that spans across the entire Chao Phraya River basin (Dang et al., 2020)—thereby including the Ping River and station P.1. Rainfall and temperature data are taken from APHRODITE (Yatagai et al., 2012) and CFSR (Saha et al., 2014), respectively. Land use data are obtained from the Global Land Cover Characterization data set (https://www.usgs.gov/centers/eros/science/), while soil data are extracted from the Harmonized World Soil Database (http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonizedworld-soil-database-v12/en/). The monthly Leaf Area Index and albedo are derived from the Terra MODIS satellite images, which represent changes in canopy and snow coverage over time. The Digital Elevation Model is created by masking the Global 30 Arc-Second Elevation DEM with the shape of the Chao Phraya River basin. What is perhaps more important here is that the VIC model represents the storage dynamics of the two largest reservoirs within the Chao Phraya basin (i.e., Bhumibol and Sirikit), but not Mae Ngat Dam, meaning that it simulates the behaviour of the Ping River (upstream of station P.1) under natural, or pre-dam development, conditions. Furthermore, the model is setup for the period 1976–2005, conveniently spanning across the pre- and post-dam development conditions (recall that the dam became operational in 1985). Further details about the model calibration and validation are available in Chowdhury et al. (2020b).

The simplest option to naturalize the streamflow time series at station P.1 is to replace the observed discharge data (over the period 1986–2005) with those simulated by VIC. However, such solution does not account for the unavoidable differences between observed and simulated data, ultimately affecting our streamflow reconstruction exercise. For this reason, we proceeded by first bias-correcting VIC's output. Specifically, we used the quantile mapping method RQUANT (Gudmundsson et al., 2012), implemented in the R package `qmap` (Gudmundsson, 2016). RQUANT has been shown to work well in bias-correcting climate models' precipitation outputs (Gudmundsson et al., 2012) and streamflow reconstruction from tree rings (Robeson et al., 2020). RQUANT first bootstraps the modeled and observed time series to obtain robust empirical quantile estimates, and then maps the quantiles of the model outputs to those of the observed daily streamflow. In our case, we calibrated the bias-correction algorithm on the pre-dam development scenario (1976–1985) and then applied it to the post-dam development one (1986–2005). The results illustrated in Figure B.5 show that the bias-corrected VIC's output is in reasonable agreement with the observed discharge.

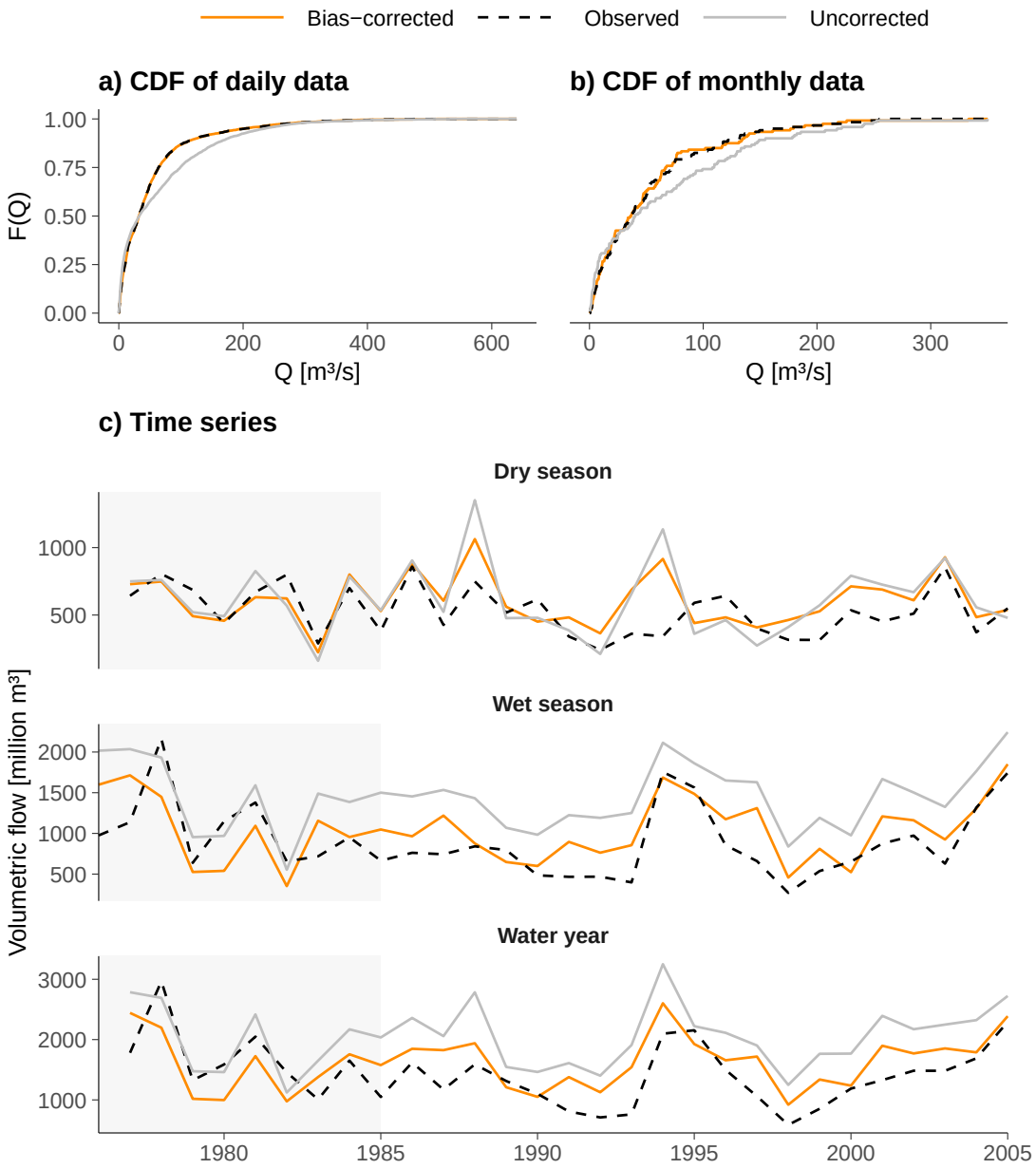

Figure B.5. Cumulative distribution function (CDF) of the daily and monthly discharge data (panels (a) and (b), respectively). The uncorrected data are simulated by the VIC hydrologic model. In panel (c), we compare the three time series (observed, simulated, and bias-corrected data) for the dry season, wet season, and water year. The pre-dam period (1976–1985) is highlighted by the shaded area.

B.2.4 Log Transformation to Improve Symmetry

We determined the distributional symmetry of the streamflow targets using the Hinkley's D metric (Hinkley, 1977). We compared the Hinkley's D of the log transformed

and non-transformed flow, and found that the log transformation improves the symmetry of distribution for all three targets (Table B.1 and Figure B.6). Therefore, we log transformed all three streamflow targets before using them in reconstruction.

Table B.1. Hinkley's D values of the streamflow targets

Target	No transformation	Log transformation
Dry season	0.14	0.02
Wet season	0.15	-0.01
Water year	0.12	-0.01

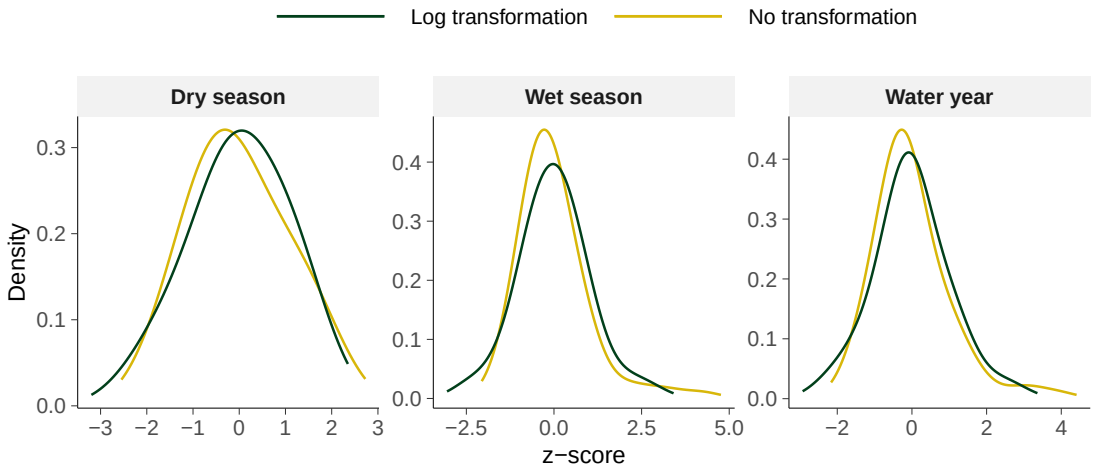

Figure B.6. Comparing densities of streamflow targets with and without log transformation. For easy comparison, the time series are converted to z-score before densities are calculated. The log transformation improves the symmetry of the distribution.

B.3 Proxy–precipitation analysis

To further explore the relationship between tree ring proxies and climate, we conducted a "climate footprint" analysis (R. Maxwell et al., 2017; Harley et al., 2017). We correlate each proxy chronology with gridded precipitation data from CRU TS4.04 (Harris et al., 2020). This is done for both monsoon (July–October) and pre-monsoon (March–May) precipitation. Correlation patterns vary remarkably among the ring width sites, and in general, ring width tend to correlate more with precipitation in the pre-monsoon months than in the monsoon months (Figures B.7 and B.8). On the other hand, $\delta^{18}O$ correlate with precipitation in the monsoon months more strongly and consistently than in the pre-monsoon months (Figure B.9). The magnitudes of correlations between $\delta^{18}O$ and precipitation also tend to be larger than that between ring width and precipitation. These results are similar to those of the proxy–streamflow correlation analysis in Figure 3.3 of the main text.

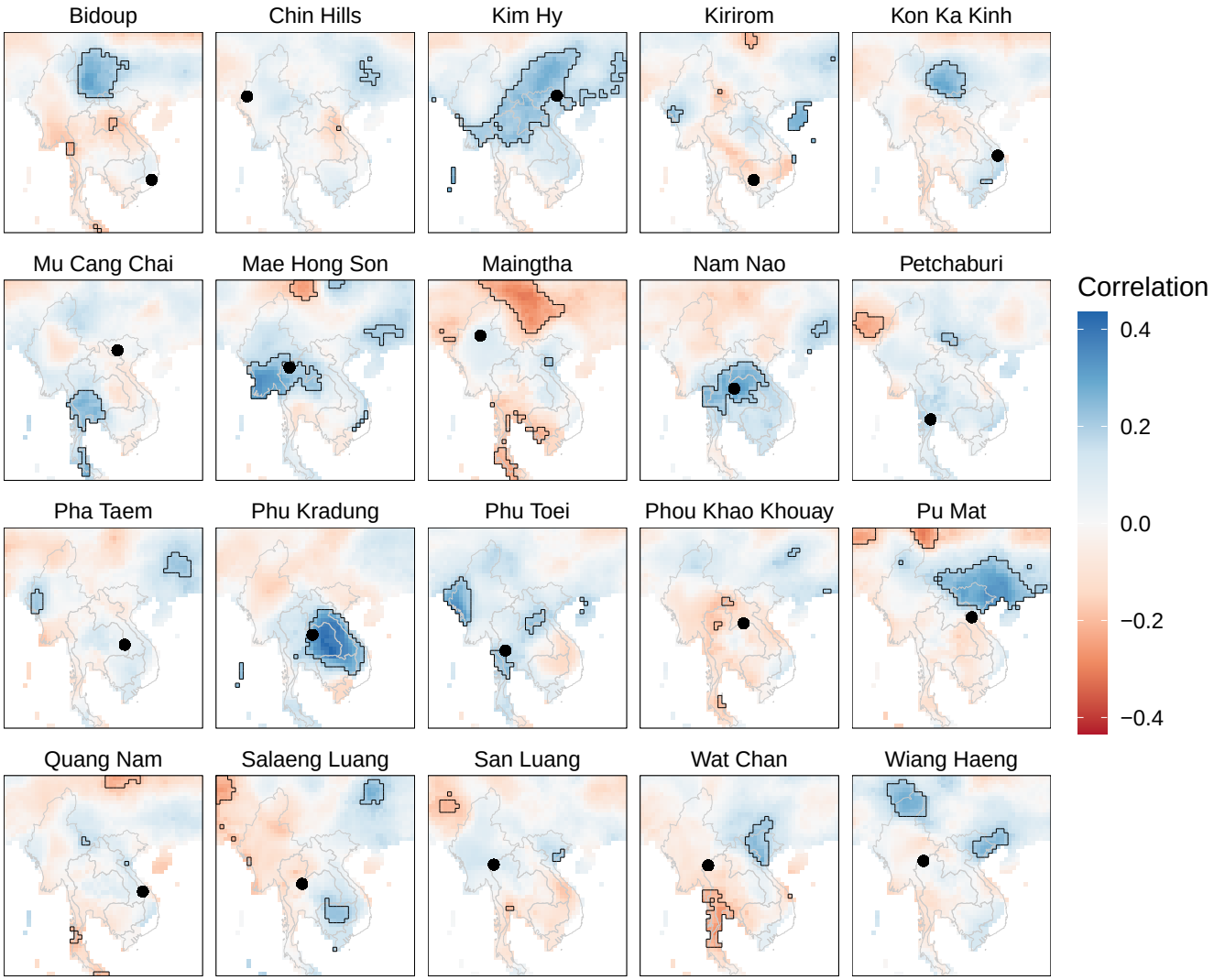

Figure B.7. Correlations between ring width chronologies (black dots) and CRU TS4.04 (Harris et al., 2020) total precipitation for July–October. The areas with significant correlation (two-tailed t-test with $\alpha = 0.05$) are enclosed in black boundaries.

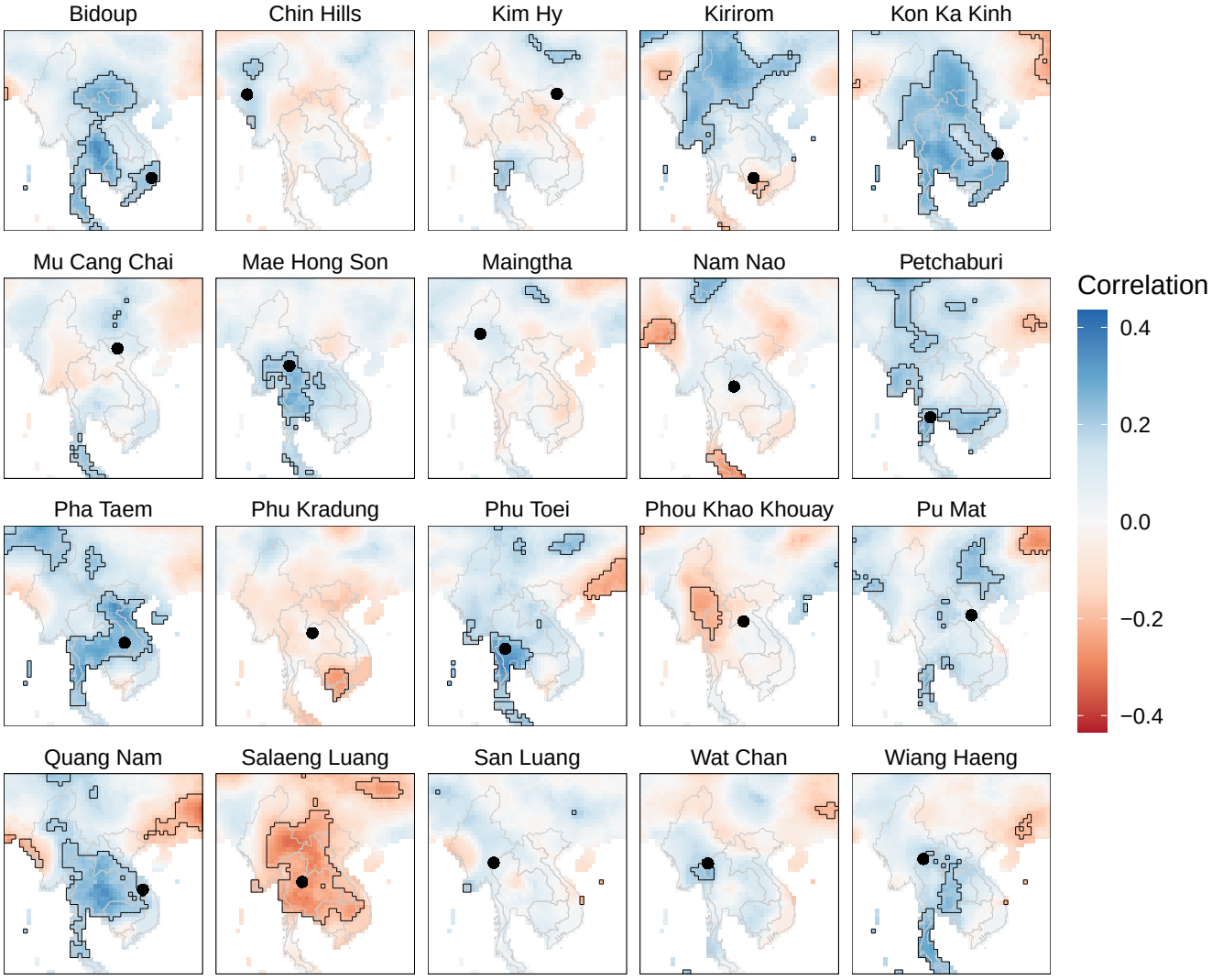

Figure B.8. Same as Figure B.7 but for March–May precipitation.

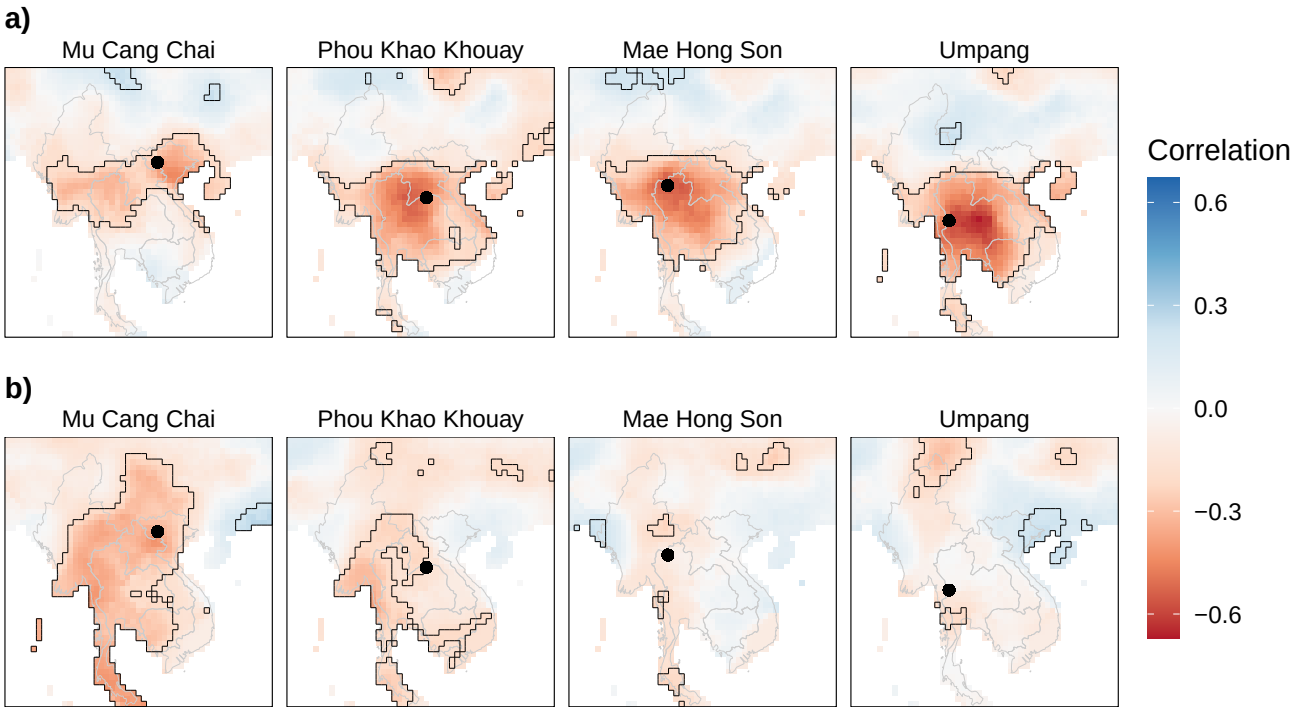

Figure B.9. Same as Figures B.7 and B.8 but the proxy used is $\delta^{18}O$. a) Correlations with monsoon precipitation (July–October). b) Correlations with pre-monsoon precipitation (March–May).

B.4 Proof of Lemma 3

Proof. Observe that J_2 is quadratic. Let

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} \mathbf{X_d} & \mathbf{X_w} & -\mathbf{X_q} \end{bmatrix} \\ \therefore \delta &= \mathbf{A}\boldsymbol{\beta} \\ \therefore J_2 &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \lambda \boldsymbol{\beta}' \mathbf{A}' \mathbf{A} \boldsymbol{\beta} \end{aligned}$$

Compute first order derivative with respect to $\boldsymbol{\beta}$ and set to zero:

$$\begin{aligned} \frac{\partial J_2}{\partial \boldsymbol{\beta}} &= -2\mathbf{X}'(\mathbf{y} - \mathbf{Xfi}) + 2\lambda \mathbf{A}'\mathbf{A}\boldsymbol{\beta} \\ &= -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + 2\lambda \mathbf{A}'\mathbf{A}\boldsymbol{\beta} \\ \frac{\partial J_2}{\partial \boldsymbol{\beta}} &= 0 \\ \implies \boldsymbol{\beta} &= (\mathbf{X}'\mathbf{X} + \lambda \mathbf{A}'\mathbf{A})^{-1}\mathbf{X}'\mathbf{y} \end{aligned}$$

■

B.5 Genetic Algorithm Implementation

In genetic algorithm (GA), each binary vector $\mathbf{p}$ is thought of as an individual that can evolve over time. Starting with an initial population, in each generation (time step), the individuals go through *mutation*, *crossover*, and *selection*. Mutation means some binary elements are switched from 0 to 1 or vice versa. Crossover between two individuals $\mathbf{p_1}$ and $\mathbf{p_2}$ means a segment of $\mathbf{p_1}$ is swapped to $\mathbf{p_2}$ and another segment of the same length is swapped from $\mathbf{p_2}$ to $\mathbf{p_1}$. Selection means only individuals with good fitness are able to live on to the next generation. Here, fitness is defined as $f(\mathbf{p})$ in Equation 3.17. As the population evolves, the fitness increases. The algorithm stops after a number of generations, and the individual with the best fitness in the final population is the optimal solution.

Two important tuning parameters in GA are the population size N and the number of generations G. In each generation, the fitness value is calculated for each individual, so $N \times G$ is the number of function evaluations (NFEs). We set $N = 500$ and $G = 500$, i.e., $NFE = 250,000$. In our experience, this value is sufficient to achieve convergence. In fact, for our experiment with P.1, the algorithm converged after about 100 generations for Model 0, and about 300 generations in Model 1 (Figure B.10). For other parameters, we used the sensible defaults set by the package `GA` (Scrucca, 2013).

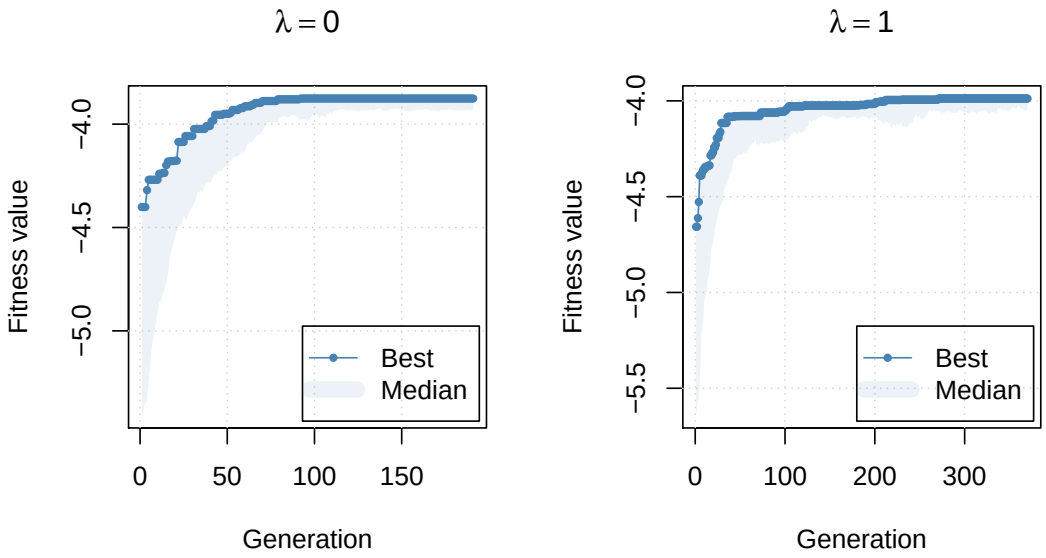

Figure B.10. Convergence of GA solutions. The shaded areas show the best 50th percentiles of all individuals, i.e., the bottom of the shaded area is the median of all individuals.

B.6 Intra-Annual Streamflow Variability

To explore the intra-annual streamflow variability that is captured in the construction, we calculated the ratio (D/Q) of dry season flow (D) to annual flow (Q) for each year, and plotted it against the total annual flow volume. We did this for both the naturalized instrumental time series and the reconstructions (Figure B.11). The years 1971, 1973, and 1975 were affected by consecutive La Niña events (see https://www.coaps.fsu.edu/jma). As a result, extremely high flow occurred in the wet season, leading to small D/Q values. As we discussed in Section 3.4.1 of the main text, these extreme high flows could not be captured by tree rings. Apart from those three extreme years, the distributions of D/Q versus Q are similar in the instrumental record and in the reconstruction. Notably, the reconstruction reveals some extreme years such as 1808 and 1815, where D/Q is more than 50%.

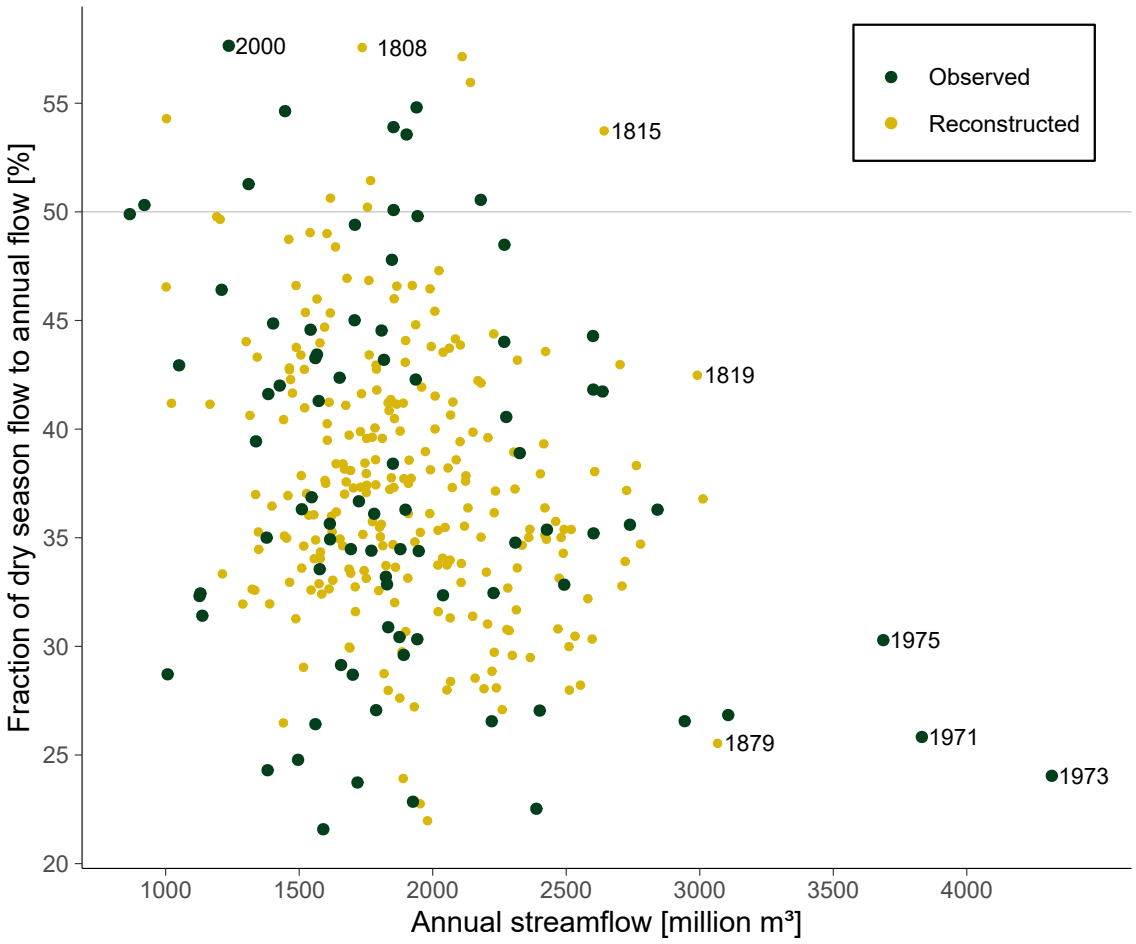

Figure B.11. Distribution of the fraction of dry season flow to total annual flow, plot against the annual flow volume. Some extreme years are annotated.

B.7 Additional Mass Balance Verifications

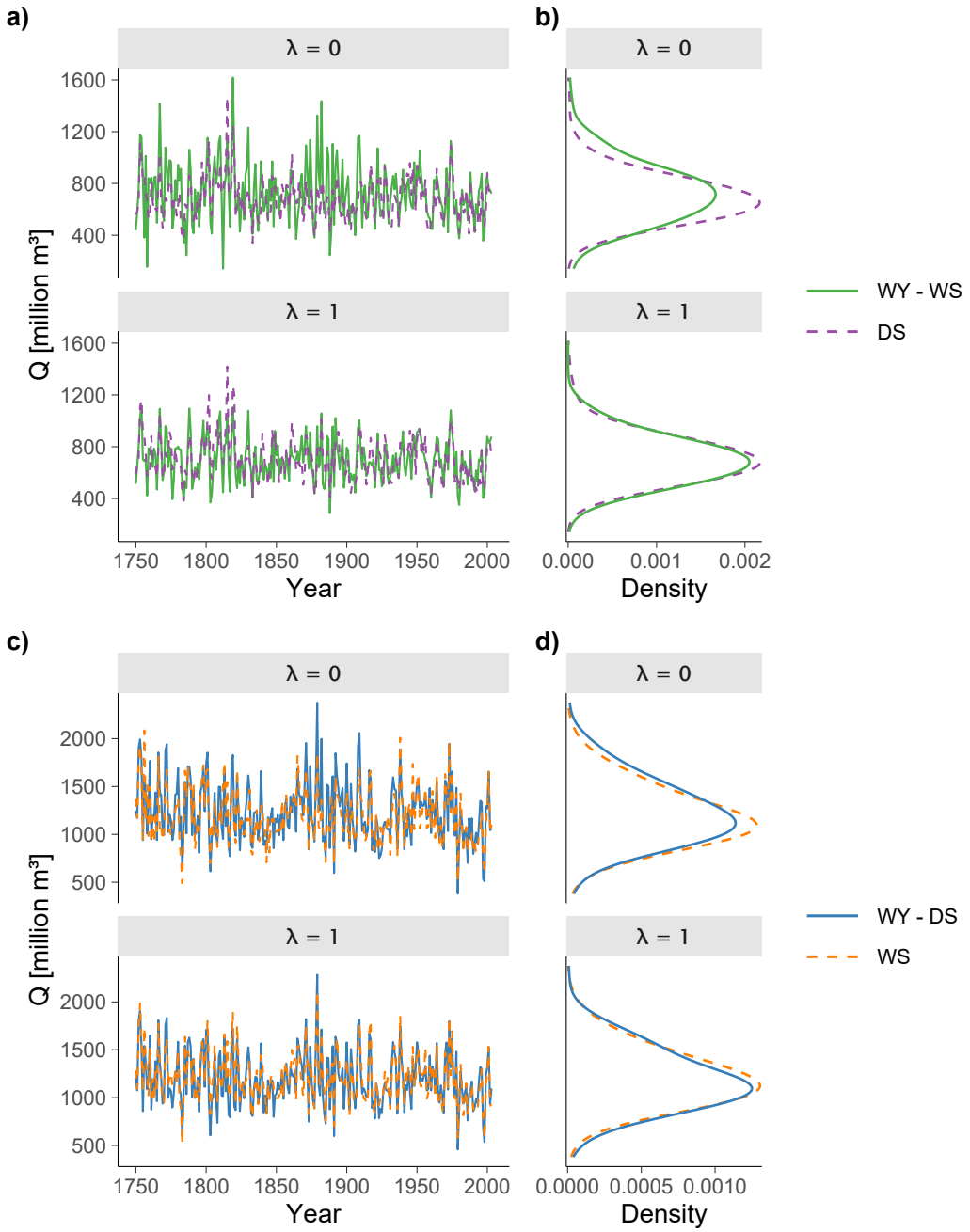

Figure B.12. Time series and distribution of the seasonal water budget, similar to Figures 3.9a and 3.9b in the main text, but for the seasons instead. DS: dry season; WS: wet season; WY: water year. Dry season's water budget can be estimated either from the dry season reconstruction directly, or by subtracting the wet season reconstruction away from the annual reconstruction. Similarly for the wet season.

B.8 Generating Sub-Annual Stochastic Streamflow

As a proof of concept, we illustrate one simple way to generate sub-annual stochastic streamflow time series from the reconstructions. We first visually examined the distribution of the dry season and wet season reconstructions. After log-transforming the time series, we plotted their univariate and bivariate distributions using kernel density estimation (Figure B.13). The marginal distributions are approximately normal; therefore, their joint distribution is also approximately normal. We then fitted a bivariate normal distribution to the time series, and generated 100 stochastic replicates of the seasonal reconstructions. Each replicate contains 254 years, sampled from the fitted bivariate normal distribution. This simple sampling scheme does not account for the autocorrelation of the time series, but it is sufficient for our illustration purposes here. The sampling process produced two ensembles of stochastic replicates for the two seasons. We then calculated the annual ensemble by summing up the individual pair of time series from the seasonal ensembles.

Figure B.14 shows a comparison between the three stochastic ensembles and the original reconstructions. The ensembles capture the variability of the original reconstructions well while also producing more extreme events (Figure B.14a). Not only that, the ensembles also faithfully reproduce the densities of the reconstructions while producing additional variability (Figure B.14b). These are desirable characteristics that will be useful in water resources applications, such as stress-testing the water system.

It should be emphasized that the annual ensemble is able to replicate the annual reconstruction well even though we did not sample from the annual distribution directly. This is only possible because the mass-balance-adjusted regression has produced nearly identical distributions for the annual water budget, both directly from the annual reconstruction and from the total seasonal reconstruction (Section 3.4.2).

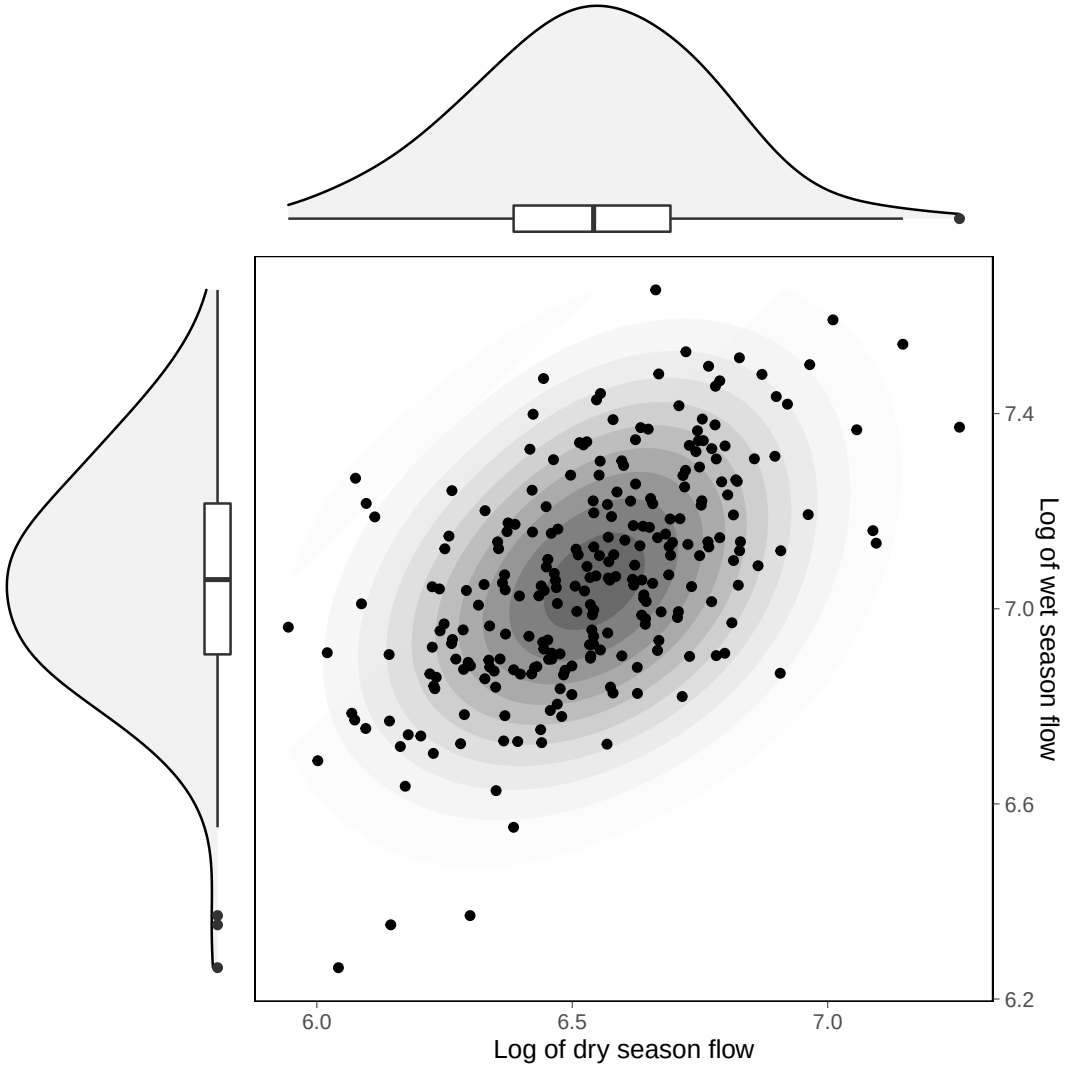

Figure B.13. Marginal and joint distributions of the dry and wet seasons' log-transformed reconstructions. Each dot represents one year. The contours underlying the dots show the bivariate probability density.

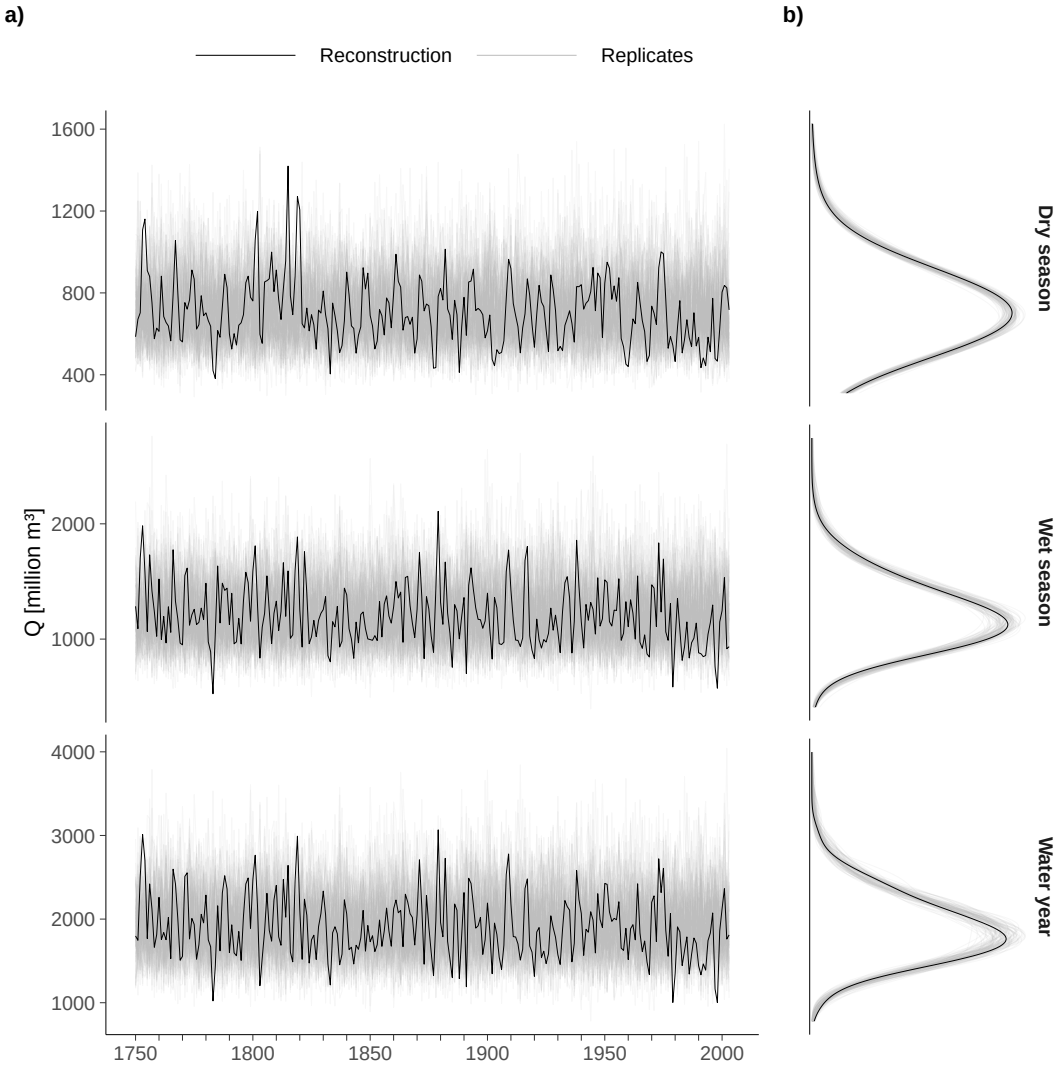

Figure B.14. Comparison between the reconstructed streamflow time series and the stochastic replicates generated from a bivariate normal distribution fitted to the reconstructions (Text B.8). a) Time series comparison. b) Density comparison.

Appendix C

Supporting Materials for Chapter 4

C.1 Previous streamflow reconstructions in Monsoon Asia

The first streamflow reconstruction in Monsoon Asia was by Davi et al. (2006). As of May 2019, 27 reconstruction studies have appeared, more than half of which were published in the last four years (Figure C.1). Each of these works studied a specific river; most of them focused on China (Table C.1).

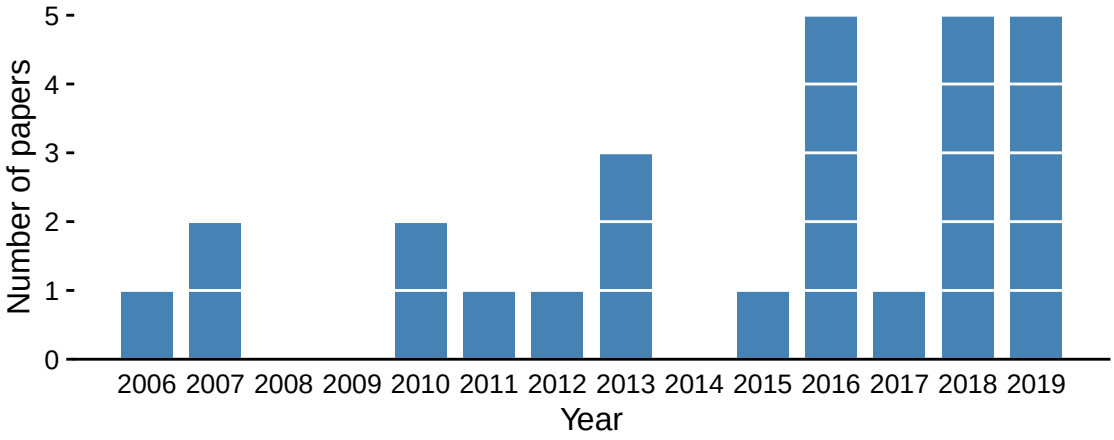

Figure C.1. Number of Monsoon Asia streamflow reconstruction papers published each year till September 2019. The publications are listed in Table C.1.

Table C.1. List of Monsoon Asia streamflow reconstruction papers

Reference	Proxy	River	Country
Davi et al. (2006)	Tree ring	Selenge	Mongolia
Yuan et al. (2007)	Tree ring	Manasi	China
X. Gou et al. (2007)	Tree ring	Yellow	China
Liu et al. (2010)	Tree ring	Heihe	China
X. H. Gou et al. (2010)	Tree ring	Yellow	China
D'Arrigo et al. (2011a)	Tree ring	Citarum	Indonesia
Yang et al. (2012)	Tree ring	Heihe	China
E. R. Cook et al. (2013)	Tree ring	Indus	Pakistan
Davi et al. (2013)	Tree ring	Kherlen	Mongolia
Pederson et al. (2013)	Tree ring	Yeruu	Mongolia
C. Xu et al. (2015)	Stalagmite δ^{18}O	Jialingjiang	China
Chen et al. (2016a)	Tree ring	Irtysh	China
Chen et al. (2016b)	Tree ring	Shiyang	China
Chen and Yuan (2016)	Tree ring	Guxiang	China
R. Zhang et al. (2016a)	Tree ring	Tuoshigan	China
R. Zhang et al. (2016b)	Tree ring	Aksu	China
Chen et al. (2017)	Tree ring	Kurshab	Kyrgyzstan
Panyushkina et al. (2018)	Tree ring	Ili	Kazakhstan
T. Zhang et al. (2018)	Tree ring	Haba	China
Rao et al. (2018)	Tree ring	Indus	Pakistan
Nguyen and Galelli (2018)	MADA[a]	Ping	Thailand
Li et al. (2018)	Tree ring	Yangtze	China
Chen et al. (2019b)	Tree ring	Lhasa	China
Chen et al. (2019a)	Tree ring	Salween	China
Yang et al. (2019)	Tree ring	Lancang	China
Li et al. (2019)	Tree ring	Yellow	China
C. Xu et al. (2019)	Tree ring δ^{18}O	Chao Phraya	Thailand

[a] Monsoon Asia Drought Atlas (E. R. Cook et al., 2010a)

C.2 Station Selection

We obtained most of our mean annual flow data from the Global Streamflow Indices and Metadata (GSIM) Archive (Do et al., 2018; Gudmundsson et al., 2018). The GSIM authors ignored missing data when calculating mean annual flow, but provided for each station the fraction of missing days for the whole record length, and the number of missing days for each year. We first selected stations with no more than 3% of missing days over the whole record length. Then, for each of these stations, we looked at each year's number of missing days, and if this number was greater than 30, we considered that year's data as missing. We adopted these criteria to avoid the situation where the mean annual flow was calculated from too many missing data. After this second step, we counted the number of non-missing years for each station, and retain only those having at least 41 years.

For the Chao Phraya, we obtained monthly flow from the Thai Royal Irrigation Department (hydro-1.net, in Thai) for stations P.1, N.1, and C.2, and calculated the mean annual flow from the monthly flow. If there were more than one month missing for any year, that year was considered missing as well (similar to what we did with the GSIM stations).

For Mekong, Yangtze, Citarum, and Brahmaputra data, we obtained annual flow directly from our colleagues, and we did not have any information on the degree of missingness. Mekong data were provided by Thanh Dang, Brahmaputra by Mukund Rao and Kamal Chowdhury, and Citarum by Rosanne D'Arrigo.

Angat River data were provided by Christoph Libisch-Lehner. The time series is almost complete, and we did not do any gap filling.

Indus River data were taken from the Supporting Information of Rao et al. (2018).

South Korea data were provided by Donghoon Lee. There were no missing data, but the longest record was only 39 years. We wanted to have a station for this country, so we made an exception for the 41-year criterion. Similarly, we made an exception to the Yeruu River (data provided by Caroline Leland): the mean annual flow here is 49.8 m^3/s, slightly less than the 50 m^3/s threshold, but we retain this record so as to have a station in Mongolia.

Figure C.2 shows the distribution of record length over the selected stations.

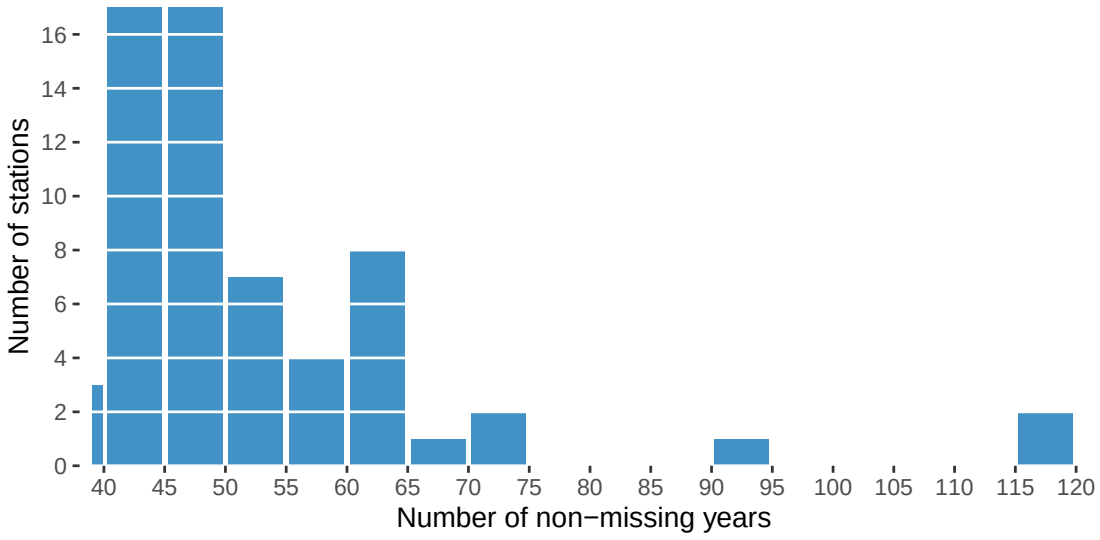

Figure C.2. Distribution of the number of non-missing years of the streamflow data set.

C.3 Streamflow data preprocessing

We determined the degree of asymmetry of the streamflow data using the Hinkley's D statistic (Hinkley, 1977), formulated according to Equation C.1:

$$D = \frac{m - \mu}{q}, \tag{C.1}$$

where m is the sample median, μ the sample mean, and q the sample inter-quartile range. If log-transforming reduces the absolute value of D for a station, then we will use the log-transformed flow as reconstruction target; otherwise we use the untransformed flow. We also check the densities of the transformed and untransformed flow visually (Figure C.3), and found that the densities are similar for most stations.

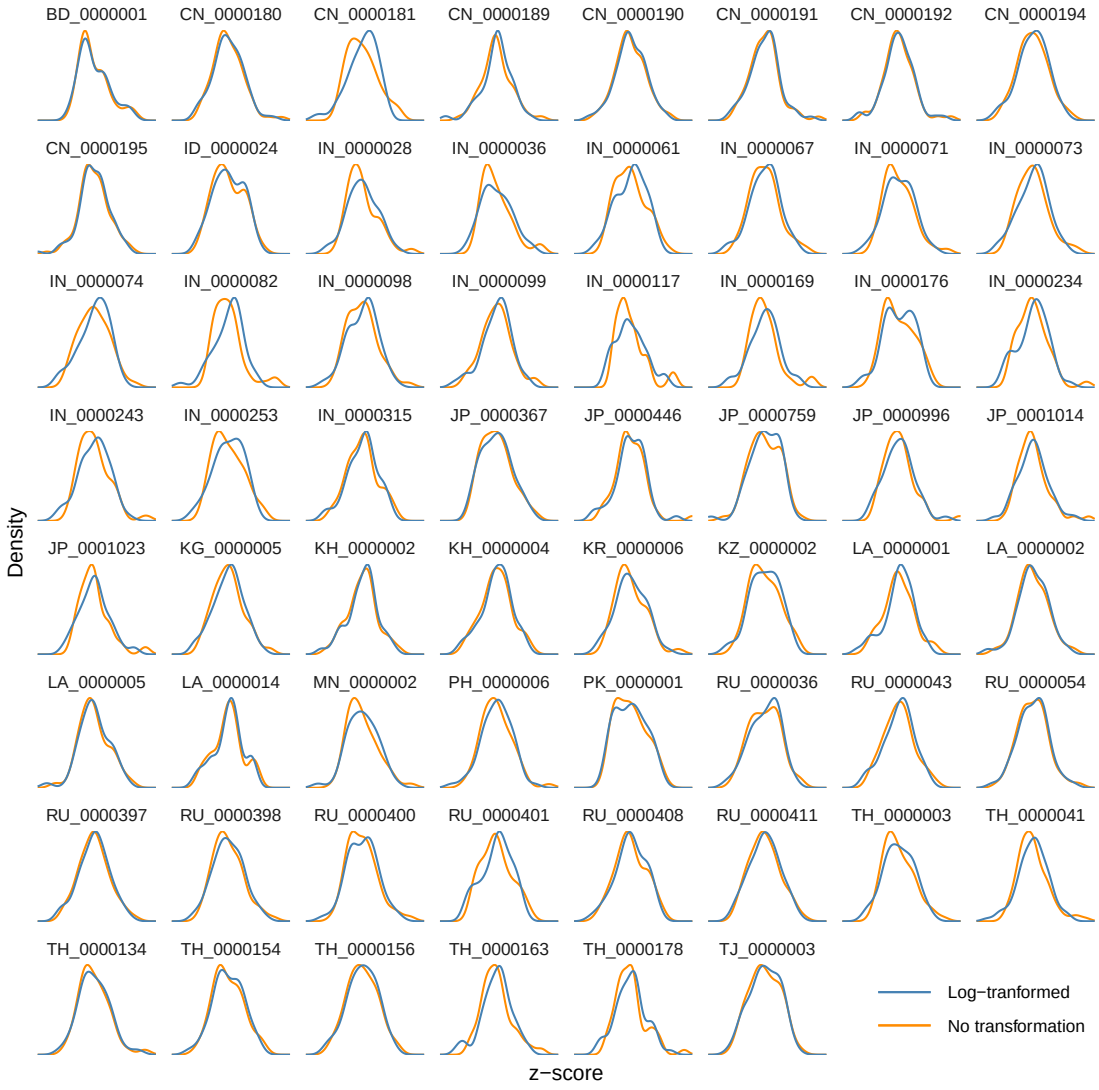

Figure C.3. Densities of the transformed and untransformed flow at each station. The densities are centralized and rescaled for comparison.

C.4 Additional Validation Exercises

C.4.1 Spatial coherence in the modern period (1950–2012)

As another reliability test for our reconstructions, we checked the spatial coherence of our reconstructions for the modern period against instrumental streamflow data, using the same standardization as in Figure 4.6. The results are shown in Figure C.4. We observe similar coherence among river basins in the instrumental data as in the reconstruction. The reconstructed streamflow also captures the extreme events very well.

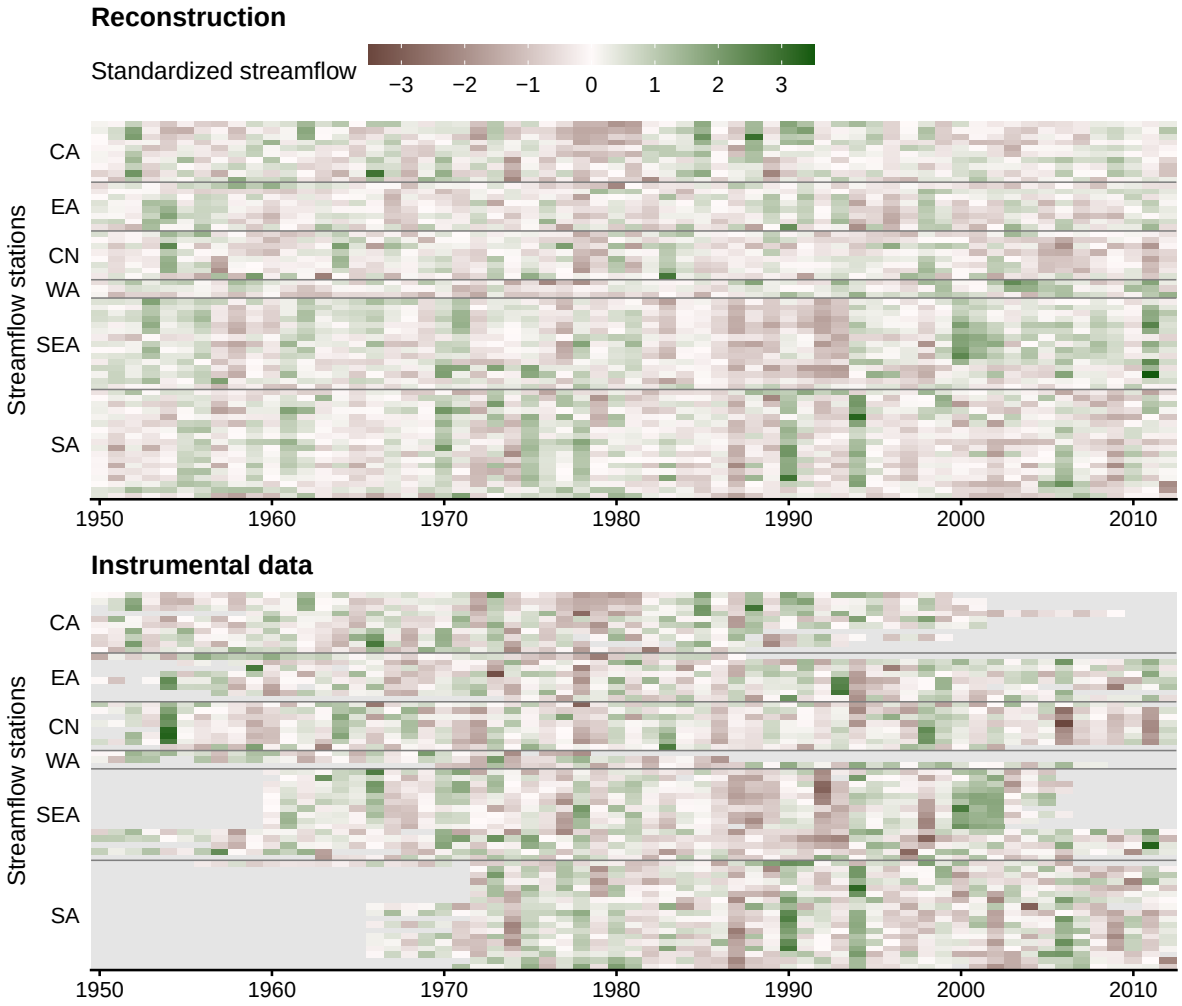

Figure C.4. Comparing the reconstructed spatiotemporal variability of streamflow in the period 1950–2012 with instrumental streamflow data. Gray areas denote no data; color scale and annotations are the same as Figure 4.6

C.4.2 Cross-correlation of all streamflow stations

The composite correlation matrix (Figure C.5 is close to symmetry about its diagonal; in other words, the reconstruction captures the correlation structure of the streamflow network.

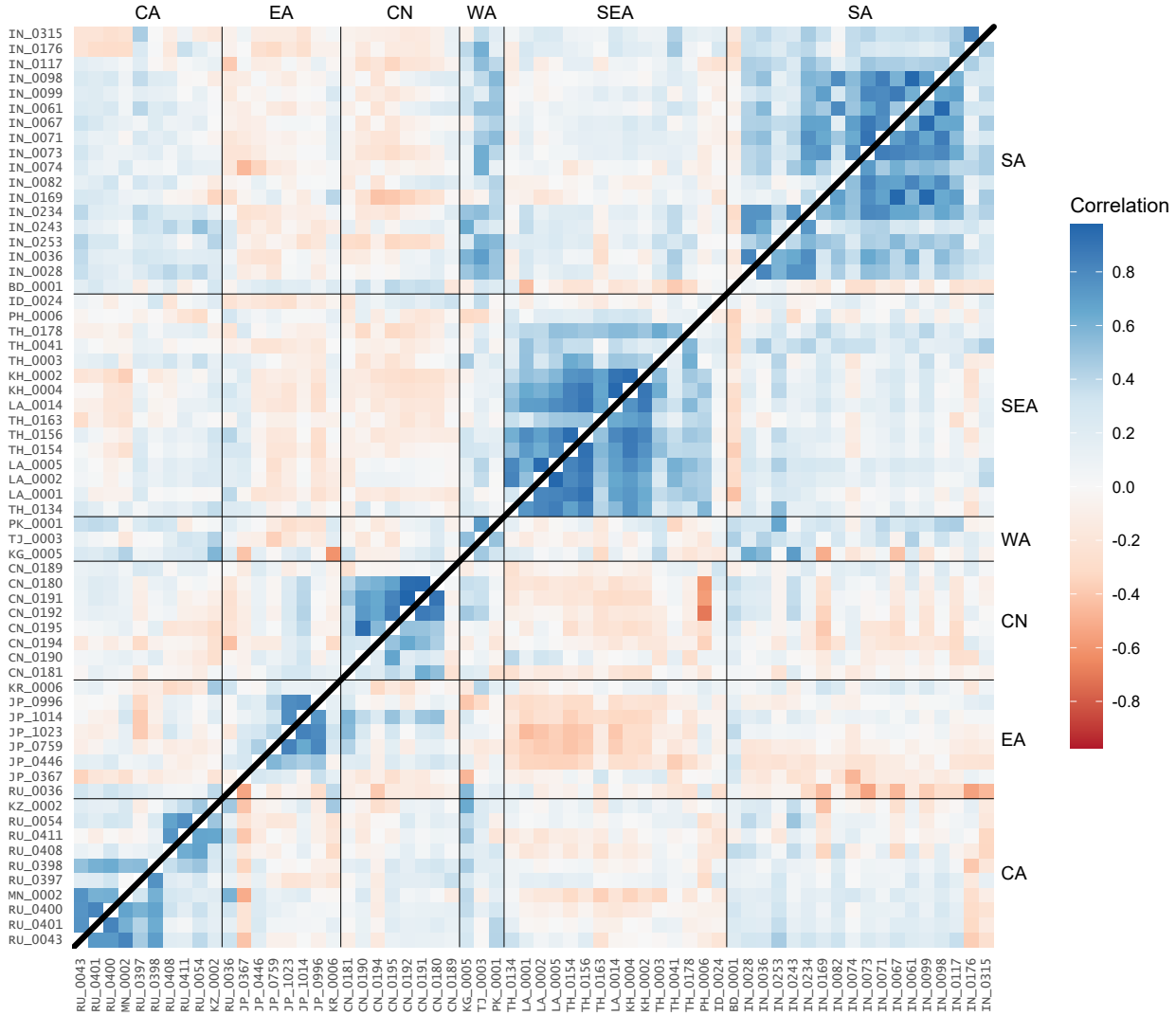

Figure C.5. Composite correlation matrix of streamflow. The top half shows the correlations in the instrumental data; the bottom half the reconstruction. Stations are grouped by their region (according to Figure 4.1) and follows the same order as in Figure 4.6.

C.4.3 Time series plots

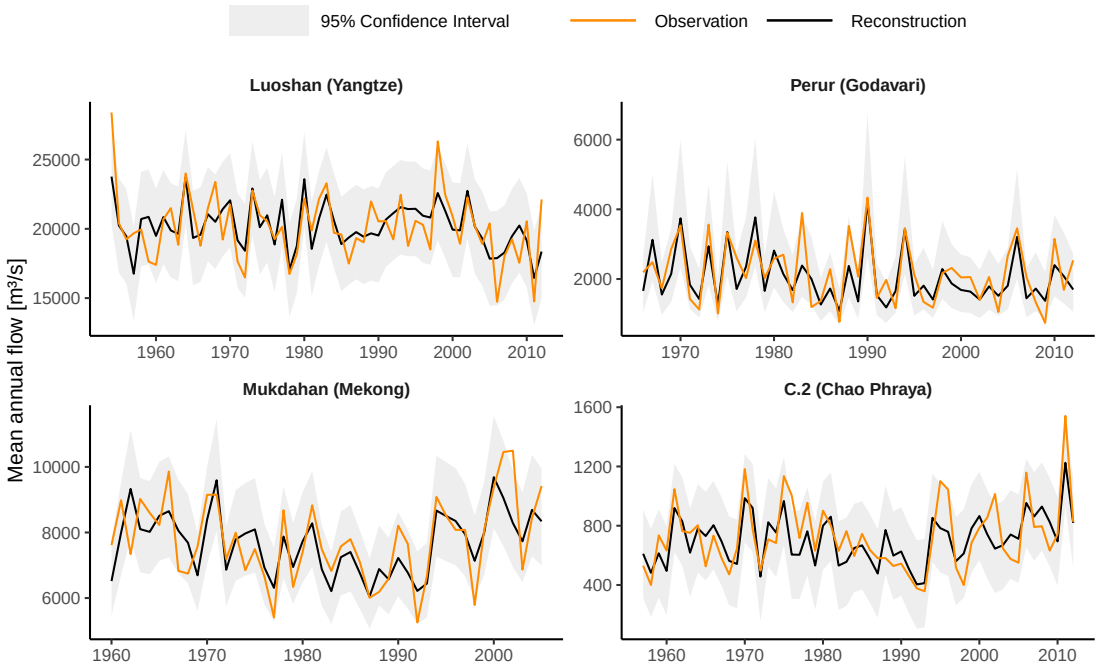

Figure C.6. Comparing reconstructions and observations for the instrumental period at four representative stations (those used in Figures 4.3 and 4.8).

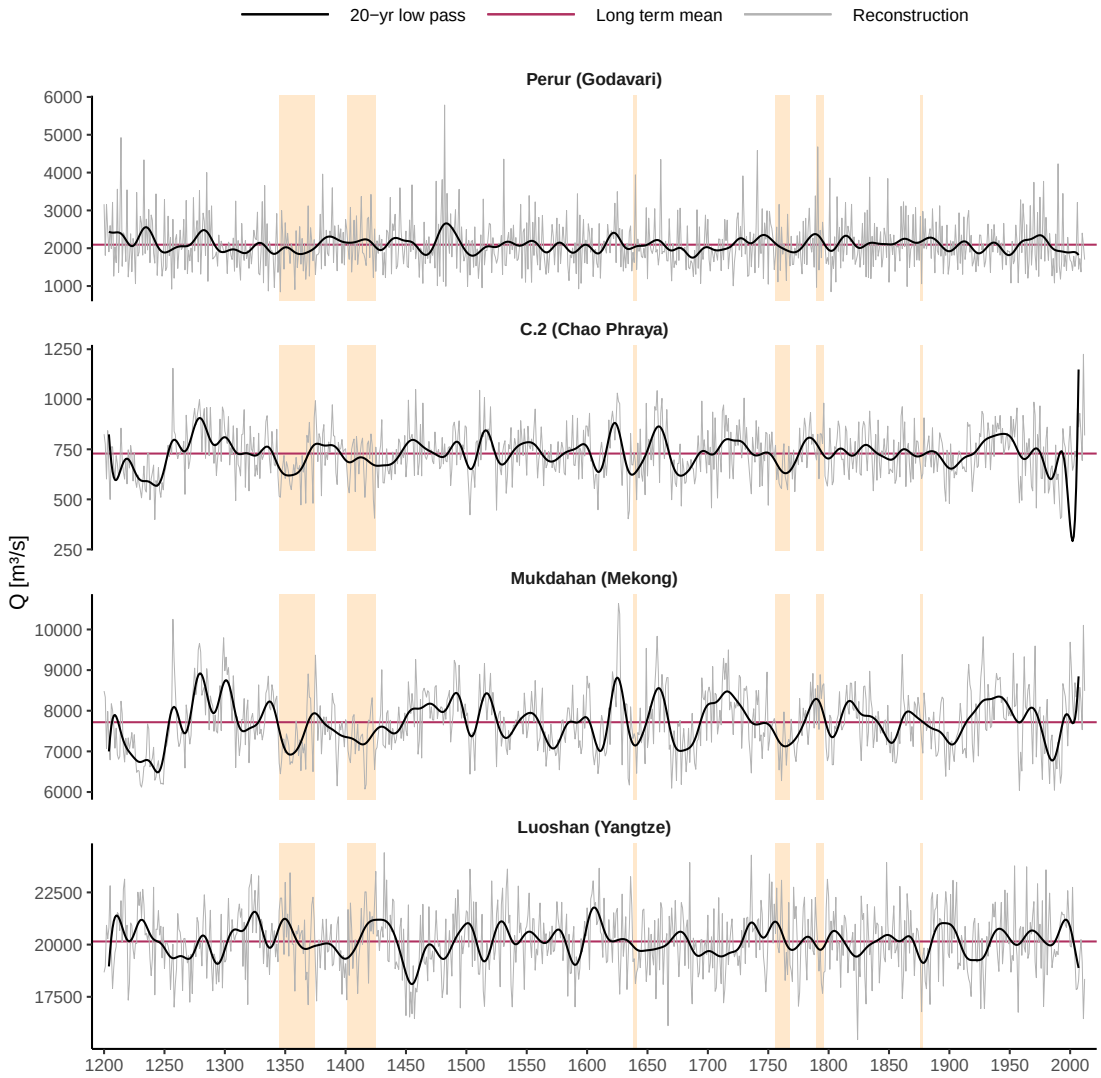

Figure C.7. Full reconstruction time series for the same four stations shown in Figure C.6. Vertical shaded areas show the megadroughts of Figure 4.6 (from left to right: Angkor Drought I, Angkor Drought II, Ming Dynasty Drought, Strange Parallels Drought, East India Drought, and Victorian Great Drought).

C.5 Additional streamflow–SST correlation analyses

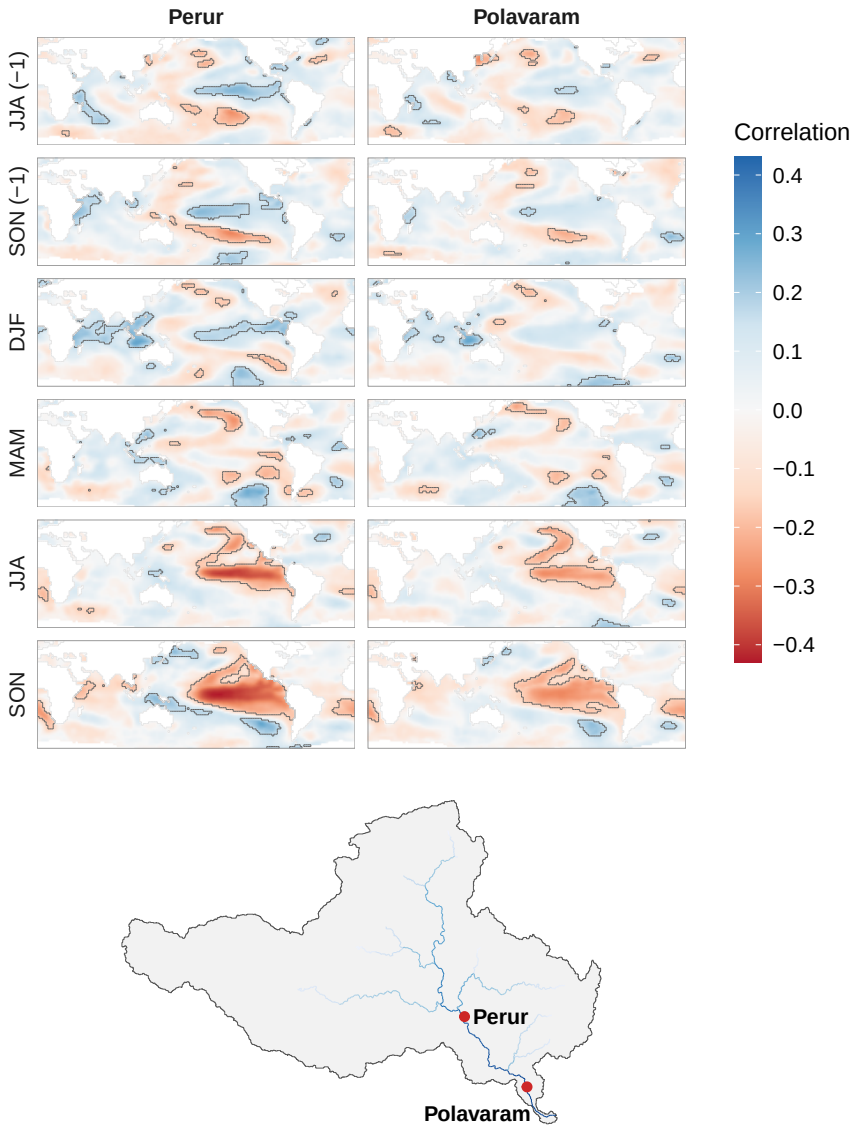

Figure C.8. Same as Figure 4.8 in the main text, but for the Godavari River.

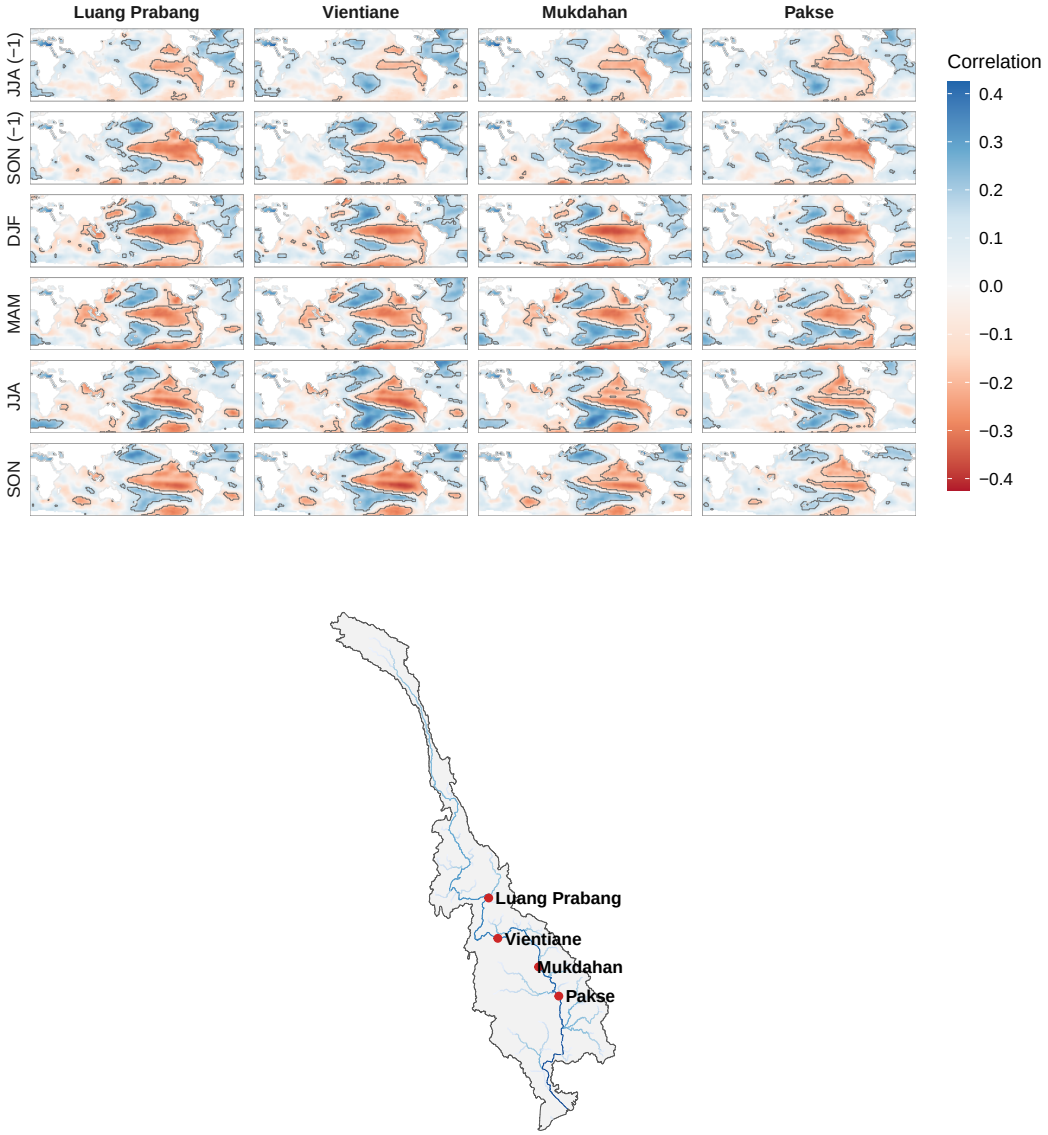

Figure C.9. Same as Figure 4.8 in the main text, but for the Mekong River.

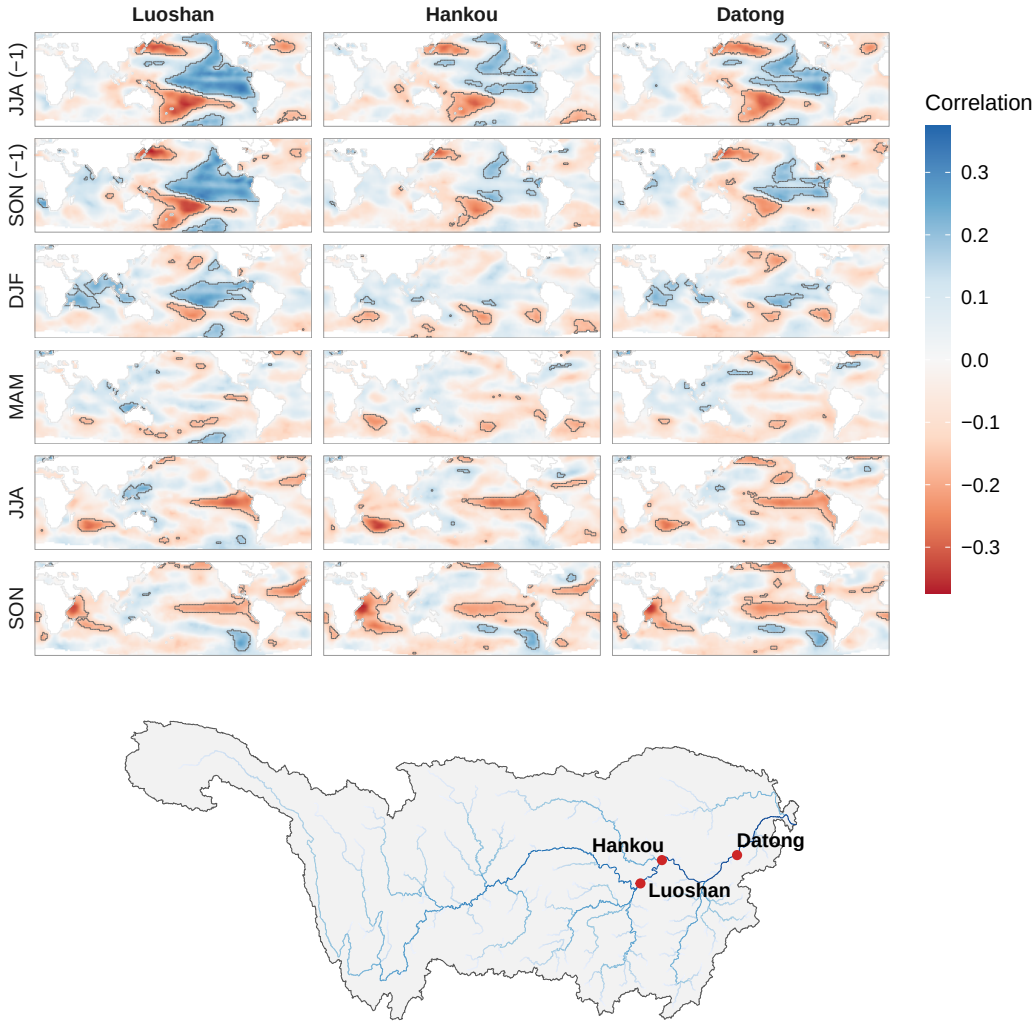

Figure C.10. Same as Figure 4.8 in the main text, but for the Yangtze River.

Appendix D

Supporting Materials for Chapter 5

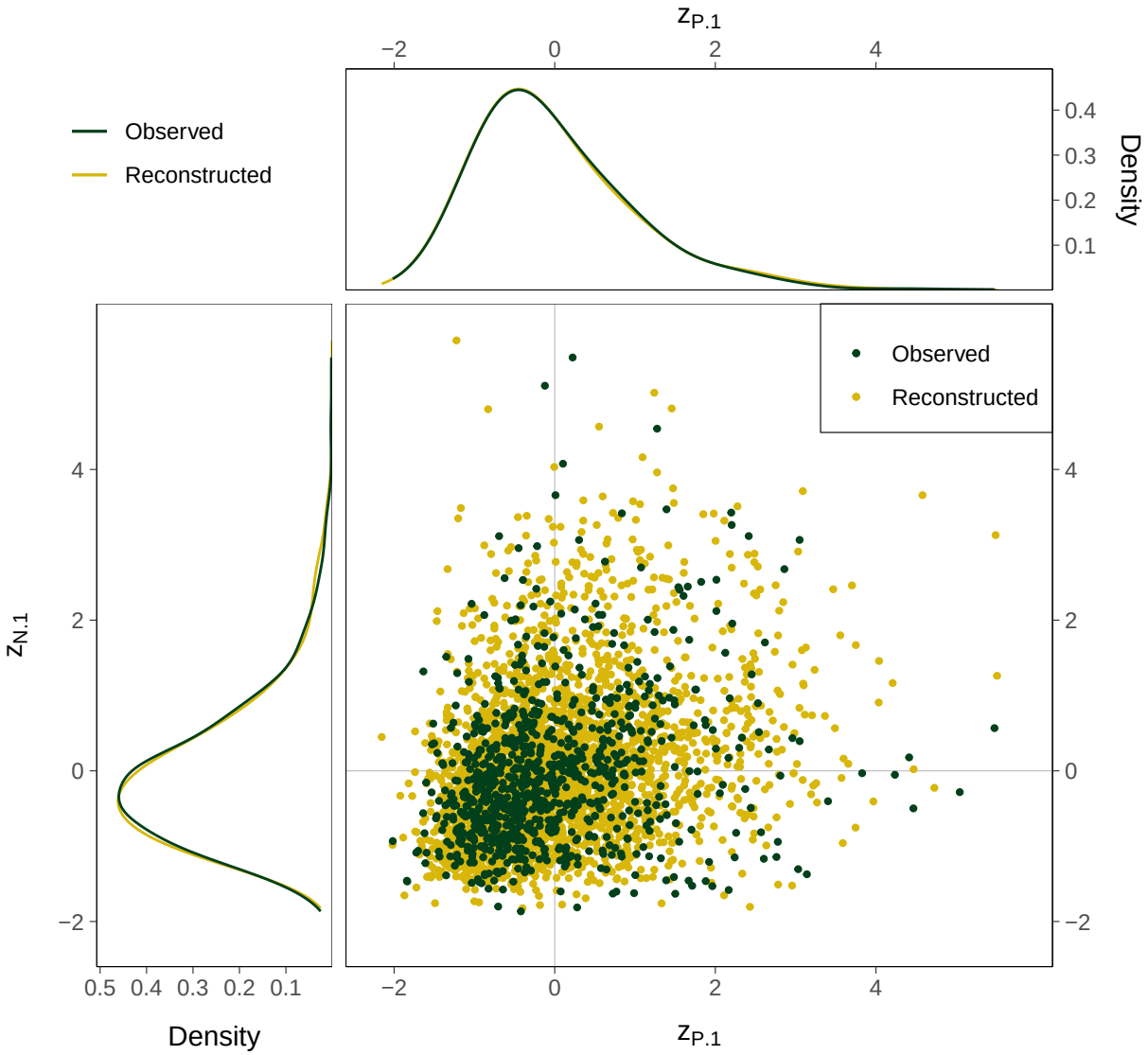

Figure D.1. Same as Figure 5.5, but using the bias-corrected reconstructions.

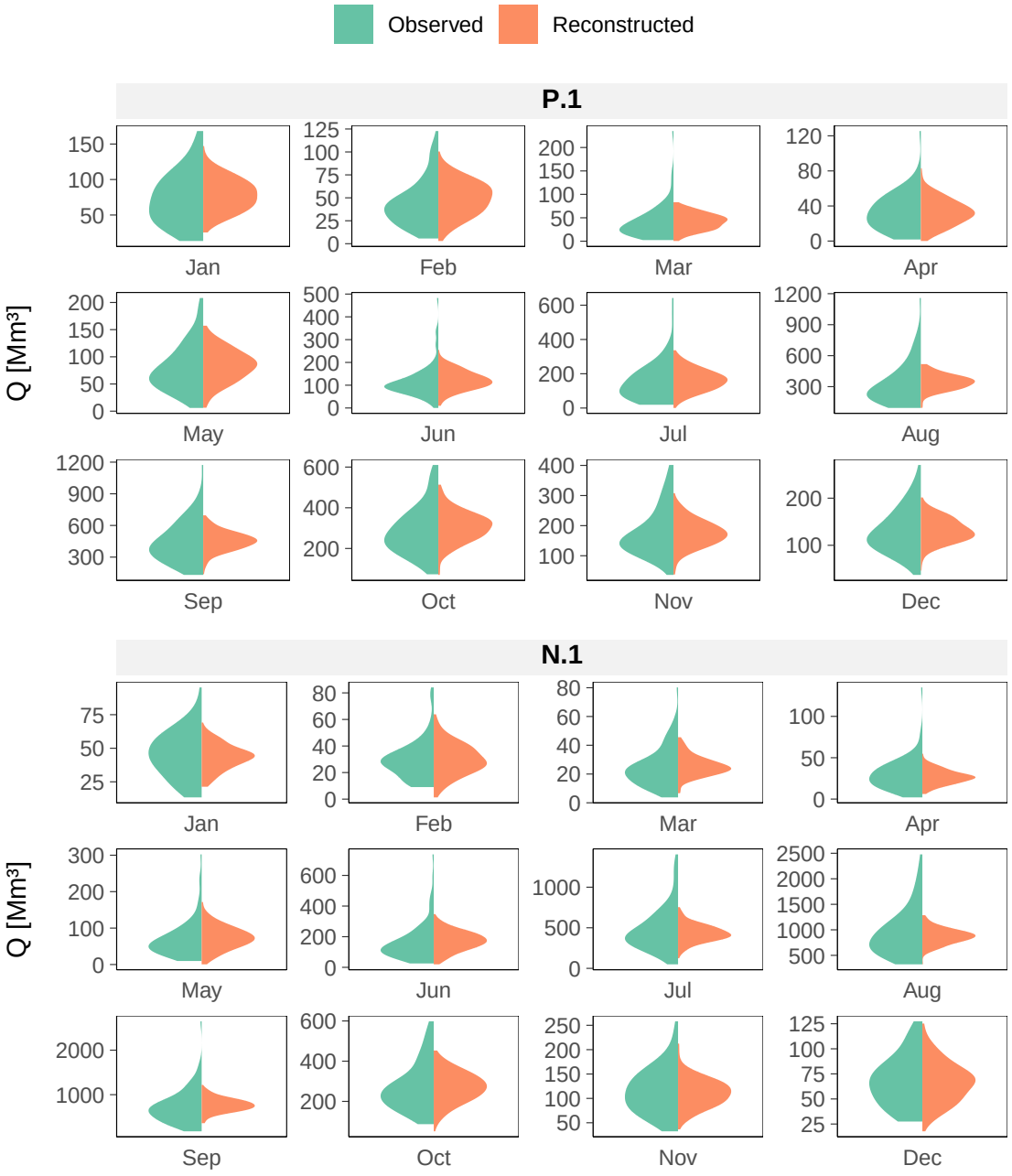

Figure D.2. Monthly streamflow distributions of the instrumental data and the reconstructions.

Bibliography

Aadhar, S., & Mishra, V. (2020). On the Projected Decline in Droughts Over South Asia in CMIP6 Multimodel Ensemble. *Journal of Geophysical Research: Atmospheres*, *125*(20). https://doi.org/10.1029/2020JD033587

Adams, K. D., Negrini, R. M., Cook, E. R., & Rajagopal, S. (2015). Annually resolved late Holocene paleohydrology of the southern Sierra Nevada and Tulare Lake, California. *Water Resources Research*, *51*(12), 9708–9724. https://doi.org/10.1002/2015WR017850

Allen, K. J., Nichols, S. C., Evans, R., Allie, S., Carson, G., Ling, F., Cook, E. R., Lee, G., & Baker, P. J. (2017). A 277 year cool season dam inflow reconstruction for Tasmania, southeastern Australia. *Water Resources Research*, *53*(1), 400–414. https://doi.org/10.1002/2016WR018906

Alley, W. M. (1984). The Palmer Drought Severity Index: Limitations and Assumptions. *Journal of Climate and Applied Meteorology*, *23*(7), 1100–1109. https://doi.org/10.1175/1520-0450(1984)023<1100:TPDSIL>2.0.CO;2

Anchukaitis, K. J., Buckley, B. M., Cook, E. R., Cook, B. I., D'Arrigo, R. D., & Ammann, C. M. (2010). Influence of volcanic eruptions on the climate of the Asian monsoon region. *Geophysical Research Letters*, *37*(22), 1–5. https://doi.org/10.1029/2010GL044843

Anchukaitis, K. J., Cook, E. R., Buckley, B. M., D'Arrigo, R. D., Cook, B. I., Jacoby, G. C., Bell, A., Penny, D., Tierney, J. E., Fletcher, R., Jungclaus, J., Ammann, C., Wilson, R. J., Ummenhofer, C., Barichivich, J., & Le, C. N. (2011). What can we learn from the Monsoon Asia Drought Atlas [MADA]?

Anderson, B. D. O., & Moore, J. B. (1979). *Optimal Filtering*. Dover Publications.

Barbarossa, V., Huijbregts, M. A., Beusen, A. H., Beck, H. E., King, H., & Schipper, A. M. (2018). FLO1K, global maps of mean, maximum and minimum annual streamflow at 1 km resolution from 1960 through 2015. *Scientific Data*, *5*(October 2017), 180052. https://doi.org/10.1038/sdata.2018.52

Bekker, M. F., Justin DeRose, R., Buckley, B. M., Kjelgren, R. K., & Gill, N. S. (2014). A 576-Year Weber River Streamflow Reconstruction from Tree Rings for Water Resource Risk Assessment in the Wasatch Front, Utah. *JAWRA Journal of the American Water Resources Association*, *50*(5), 1338–1348. https://doi.org/10.1111/jawr.12191

Best, J. (2019). Anthropogenic stresses on the world's big rivers. *Nature Geoscience*, *12*(1), 7–21. https://doi.org/10.1038/s41561-018-0262-x

Boers, N., Goswami, B., Rheinwalt, A., Bookhagen, B., Hoskins, B., & Kurths, J. (2019). Complex networks reveal global pattern of extreme-rainfall teleconnections. *Nature*. https://doi.org/10.1038/s41586-018-0872-x

Borgomeo, E., Farmer, C. L., & Hall, J. W. (2015a). 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. (2015b). 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

Bracken, C., Rajagopalan, B., & Woodhouse, C. A. (2016). A Bayesian hierarchical nonhomogeneous hidden Markov model for multisite streamflow reconstructions. *Water Resources Research*, *52*(10), 7837–7850. https://doi.org/10.1002/2016WR018887

Briffa, K. R., Jones, P. D., Pilcher, J. R., & Hughes, M. K. (1988). Reconstructing Summer Temperatures in Northern Fennoscandinavia Back to A.D. 1700 Using Tree-Ring Data from Scots Pine. *Arctic and Alpine Research*, *20*(4), 385. https://doi.org/10.2307/1551336

Briffa, K. R., Jones', P. D., Schweingruber, F. H., & Osborn, T. J. (1998). Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years. *Nature*, *393*(6684), 450–455. https://doi.org/10.1038/30943

Brown, C. (2010). The End of Reliability. *Journal of Water Resources Planning and Management*, *136*(2), 143–145. https://doi.org/10.1061/(asce)wr.1943-5452.65

Buckley, B. M., Anchukaitis, K. J., Penny, D., Fletcher, R., Cook, E. R., Sano, M., Nam, L. C., Wichienkeeo, A., Minh, T. T., & Hong, T. M. (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., Arrigo, R. D., 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. (2007a). Analyses of growth rings of Pinus merkusii from Lao P.D.R. *Forest Ecology and Management*, *253*(1-3), 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., Stahle, D. K., Davi, N. K., Nguyen, T. Q. T., Le, C. N., & Wilson, R. J. (2018). Blue intensity from a tropical conifer's annual rings for climate reconstruction: An ecophysiological perspective. *Dendrochronologia*, *50*(May), 10–22. https://doi.org/10.1016/j.dendro.2018.04.003

Buckley, B. M., Palakit, K., Duangsathaporn, K., Sanguantham, P., & Prasomsin, P. (2007b). 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., Le, C. N., Ton, T. M., Bui, T. H., & Nguyen, V. T. (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

Chen, F., He, Q., Bakytbek, E., Yu, S., & Zhang, R. (2017). Reconstruction of a long streamflow record using tree rings in the upper Kurshab River (Pamir-Alai Mountains) and its application to water resources management. *International Journal of Water Resources Development*, *33*(6), 976–986. https://doi.org/10.1080/07900627.2016.1238347

Chen, F., Shang, H., Panyushkina, I., Meko, D. M., Li, J., Yuan, Y., Yu, S., Chen, F., He, D., & Luo, X. (2019a). 500-year tree-ring reconstruction of Salween River streamflow related to the history of water supply in Southeast Asia. *Climate Dynamics*, (0123456789). https://doi.org/10.1007/s00382-019-04948-1

Chen, F., Shang, H., Panyushkina, I. P., Meko, D. M., Yu, S., Yuan, Y., & Chen, F. (2019b). Tree-ring reconstruction of Lhasa River streamflow reveals 472 years of hydrologic change on southern Tibetan Plateau. *Journal of Hydrology*, *572*, 169–178. https://doi.org/10.1016/j.jhydrol.2019.02.054

Chen, F., & Yuan, Y.-j. (2016). Streamflow reconstruction for the Guxiang River, eastern Tien Shan (China): linkages to the surrounding rivers of Central Asia. *Environmental Earth Sciences*, *75*(13), 1049. https://doi.org/10.1007/s12665-016-5849-1

Chen, F., Yuan, Y., Davi, N. K., & Zhang, T. (2016a). Upper Irtysh River flow since AD 1500 as reconstructed by tree rings, reveals the hydroclimatic signal of inner Asia. *Climatic Change*, *139*(3-4), 651–665. https://doi.org/10.1007/s10584-016-1814-y

Chen, F., Yuan, Y.-j., Zhang, R.-b., Wang, H.-q., Shang, H.-m., Zhang, T.-w., Qin, L., & Fan, Z.-a. (2016b). Shiyang River streamflow since AD 1765, reconstructed by tree rings, contains far-reaching hydro-climatic signals over and beyond the mid-latitude Asian continent. *Hydrological Processes*, *30*(13), 2211–2222. https://doi.org/10.1002/hyp.10788

Cheng, S., & Sabes, P. N. (2006). Modeling Sensorimotor Learning with Linear Dynamical Systems. *Neural Computation*, *18*(4), 760–793. https://doi.org/10.1162/089976606775774651

Chowdhury, A. F. M. K., Dang, T. D., Bagchi, A., & Galelli, S. (2020a). Expected Benefits of Laos' Hydropower Development Curbed by Hydroclimatic Variability and Limited Transmission Capacity: Opportunities to Reform. *Journal of Water Resources Planning and Management*, *146*(10), 05020019. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001279

Chowdhury, A. F. M. K., Dang, T. D., Nguyen, H. T. T., Koh, R., & Galelli, S. (2020b). The Greater Mekong's climate-water-energy nexus: how ENSO-triggered regional droughts affect power supply and CO2 emissions. *Earth and Space Science Open Archive*. https://doi.org/10.1002/essoar.10504393.1

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., & Kairiukstis, L. A. (Eds.). (1990). *Methods of Dendrochronology*. Springer Netherlands. https://doi.org/10.1007/978-94-015-7879-0

Cook, E. R. (2015). Developing MADAv2 Using The Point-By-Point Regression Climate Field Reconstruction Method.

Cook, E. R., Anchukaitis, K. J., Buckley, B. M., D'Arrigo, R. D., Jacoby, G. C., & Wright, W. E. (2010a). Asian Monsoon Failure and Megadrought During the Last Millennium. *Science*, *328*(5977), 486–489. https://doi.org/10.1126/science.1185188

Cook, E. R., Meko, D. M., Stahle, D. W., & Cleaveland, M. K. (1999). Drought Reconstructuions for the continental United States. *Journal of Climate*, *12*(4), 1145–1162. https://doi.org/10.1175/1520-0442(1999)012<1145:DRFTCU>2.0.CO;2

Cook, E. R., Palmer, J. G., Ahmed, M., Woodhouse, C. A., Fenwick, P., Zafar, M. U., Wahab, M., & Khan, N. (2013). Five centuries of Upper Indus River flow from tree rings. *Journal of Hydrology*, *486*(August 2018), 365–375. https://doi.org/10.1016/j.jhydrol.2013.02.004

Cook, E. R., Seager, R., Heim, R. R., Vose, R. S., Herweijer, C., & Woodhouse, C. A. (2010b). Megadroughts in North America: placing IPCC projections of hydroclimatic change in a long-term palaeoclimate context. *Journal of Quaternary Science*, *25*(1), 48–61. https://doi.org/10.1002/jqs.1303

Cook, E. R., Seager, R., Kushnir, Y., Briffa, K. R., Büntgen, U., Frank, D., Krusic, P. J., Tegel, W., Schrier, G. V., Andreu-Hayles, L., Baillie, M., Baittinger, C., Bleicher, N., Bonde, N., Brown, D., Carrer, M., Cooper, R., Eùfar, K., DIttmar, C., … Zang, C. S. (2015). Old World megadroughts and pluvials during the Common Era. *Science Advances*, *1*(10), 1–10. https://doi.org/10.1126/sciadv.1500561

Cook, E. R., Solomina, O., Matskovsky, V., Cook, B. I., Agafonov, L., Berdnikova, A., Dolgova, E., Karpukhin, A., Knysh, N., Kulakova, M., Kuznetsova, V., Kyncl, T., Kyncl, J., Maximova, O., Panyushkina, I., Seim, A., Tishin, D., Ważny, T., & Yermokhin, M. (2020). The European Russia Drought Atlas (1400–2016 CE). *Climate Dynamics*, *54*(3-4), 2317–2335. https://doi.org/10.1007/s00382-019-05115-2

Cooper, D. M., & Wood, E. F. (1982a). Identification of multivariate time series and multivariate input-output models. *Water Resources Research*, *18*(4), 937–946. https://doi.org/10.1029/WR018i004p00937

Cooper, D. M., & Wood, E. F. (1982b). Parameter estimation of multiple input-output time series models: Application to rainfall-runoff processes. *Water Resources Research*, *18*(5), 1352–1364. https://doi.org/10.1029/WR018i005p01352

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

Dai, A., Trenberth, K. E., & Qian, T. (2004). A Global Dataset of Palmer Drought Severity Index for 1870–2002: Relationship with Soil Moisture and Effects of Surface Warming. *Journal of Hydrometeorology*, *5*(6), 1117–1130. https://doi.org/10.1175/JHM-386.1

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

D'Arrigo, R., Abram, N. J., Ummenhofer, C., Palmer, J. G., & Mudelsee, M. (2011a). 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. (2011b). Three centuries of Myanmar monsoon climate variability inferred from teak tree rings. *Geophysical Research Letters*, *38*(24), 1–5. https://doi.org/10.1029/2011GL049927

Davi, N. K., Jacoby, G. C., Curtis, A. E., & Baatarbileg, N. (2006). Extension of drought records for central Asia using tree rings: West-central Mongolia. *Journal of Climate*, *19*(1), 288–299. https://doi.org/10.1175/JCLI3621.1

Davi, N. K., Pederson, N., Leland, C., Nachin, B., Suran, B., & Jacoby, G. C. (2013). Is eastern Mongolia drying? A long-term perspective of a multidecadal trend. *Water Resources Research*, *49*(1), 151–158. https://doi.org/10.1029/2012WR011834

DeGroot, M. H., & Schervish, M. J. (2012). *Probability and Statistics*. Addison-Wesley.

Delgado, J. M., Merz, B., & Apel, H. (2012). A climate-flood link for the lower Mekong River. *Hydrology and Earth System Sciences*, *16*(5), 1533–1541. https://doi.org/10.5194/hess-16-1533-2012

Dempster, A., Laird, N., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B Methodological*, *39*(1), 1–38.

DeRose, R., Bekker, M. F., Wang, S. Y. S., Buckley, B. M., Kjelgren, R., Bardsley, T., Rittenour, T., & Allen, E. (2015). A millennium-length reconstruction of Bear River stream flow, Utah. *Journal of Hydrology*, *529*(P2), 524–534. https://doi.org/10.1016/j.jhydrol.2015.01.014

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

Do, H. X., Gudmundsson, L., Leonard, M., & Westra, S. (2018). The Global Streamflow Indices and Metadata Archive (GSIM) – Part 1: The production of a daily streamflow archive and metadata. *Earth System Science Data*, *10*(2), 765–785. https://doi.org/10.5194/essd-10-765-2018

Douglass, A. E. (1914). A method of estimating rainfall by the growth of trees. *Bulletin of the American Geographical Society*, *46*(5), 321–336.

Douglass, A. E. (1919). Climatic Cycles and Tree-growth. Volume I: A study of the annual rings of trees in relation to climate and solar activity. *Publication no. 289*. Carnegie Institution of Washington.

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

Dunbar, B., Wellington, M., & Colgan, W. (1994). Eastern Pacific sea surface temperature since 1600 A.D.: The record of climate variability in Galápagos corals. *Paleobiology*, *9*(2), 291–315. https://doi.org/10.1029/93PA03501

Folland, C. K., Parker, D. E., Colman, A. W., & Washington, R. (1999). Large Scale Modes of Ocean Surface Temperature Since the Late Nineteenth Century. *Beyond el niño* (pp. 73–102). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-58369-8_4

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 and 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. (n.d.). Time to use dendrohydrological data in water resources management? *(submitted to Journal of Water Resources Planning and Management)*.

Gallant, A. J. E., & Gergis, J. (2011). An experimental streamflow reconstruction for the River Murray, Australia, 1783-1988. *Water Resources Research*, *47*(4), W00G04. https://doi.org/10.1029/2010WR009832

Gangopadhyay, S., McCabe, G. J., & Woodhouse, C. A. (2015). Beyond annual streamflow reconstructions for the Upper Colorado River Basin: A paleo-water-balance approach. *Water Resources Research*, *51*(12), 9763–9774. https://doi.org/10.1002/2015WR017283

Gao, C., Robock, A., Self, S., Witter, J. B., Steffenson, J. P., Clausen, H. B., Siggaard-Andersen, M. L., Johnsen, S., Mayewski, P. A., & Ammann, C. (2006). The 1452 or 1453 A.D. Kuwae eruption signal derived from multiple ice core records: Greatest volcanic sulfate event of the past 700 years. *Journal of Geophysical Research Atmospheres*, *111*(12), 1–11. https://doi.org/10.1029/2005JD006710

Genuer, R., Poggi, J.-m., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, *31*(14), 2225–2236. https://doi.org/10.1016/j.patrec.2010.03.014

Ghahramani, Z., & Hinton, G. E. (1996). Parameter estimation for linear dynamical systems.

Goodkin, N. F., Bolton, A., Hughen, K. A., Karnauskas, K. B., Griffin, S., Phan, K. H., Vo, S. T., Ong, M. R., & Druffel, E. R. (2019). East Asian Monsoon Variability Since the Sixteenth Century. *Geophysical Research Letters*, *46*(9), 4790–4798. https://doi.org/10.1029/2019GL081939

Gou, X. H., Deng, Y., Chen, F. H., Yang, M. X., Fang, K. Y., Gao, L. L., Yang, T., & Zhang, F. (2010). Tree ring based streamflow reconstruction for the Upper Yellow River over the past 1234 years. *Chinese Science Bulletin*, *55*(36), 4179–4186. https://doi.org/10.1007/s11434-010-4215-z

Gou, X., Chen, F., Cook, E. R., Jacoby, G. C., 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*(6), W06434. https://doi.org/10.1029/2006WR005705

Graham, N. E., & Hughes, M. K. (2007). Reconstructing the Mediaeval low stands of Mono Lake, Sierra Nevada, California, USA. *The Holocene*, *17*(8), 1197–1210. https://doi.org/10.1177/0959683607085126

Groslier, B. P. (1979). La cité hydraulique angkorienne: exploitation ou surexploitation du sol ? *Bulletin de l'Ecole française d'Extrême-Orient*, *66*(1), 161–202. https://doi.org/10.3406/befeo.1979.4014

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

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

Gudmundsson, L., Do, H. X., Leonard, M., & Westra, S. (2018). The Global Streamflow Indices and Metadata Archive (GSIM) – Part 2: Quality control, time-series indices and homogeneity assessment. *Earth System Science Data*, *10*(2), 787–804. https://doi.org/10.5194/essd-10-787-2018

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

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1-2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hansen, K. G., Buckley, B. M., Zottoli, B., D'Arrigo, R. D., Nam, L. C., Van Truong, V., Nguyen, D. T., & Nguyen, H. X. (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. *Nevada Agricultural Experiment Station Bulletin No. 141*.

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

Hart, S. J., Smith, D. J., & Clague, J. J. (2010). A multi-species dendroclimatic reconstruction of Chilko River streamflow, British Columbia, Canada. *Hydrological Processes*, *24*(19), 2752–2761. https://doi.org/10.1002/hyp.7674

Henley, B. J. (2017). Pacific decadal climate variability: Indices, patterns and tropical-extratropical interactions. *Global and Planetary Change*, *155*(October 2016), 42–55. https://doi.org/10.1016/j.gloplacha.2017.06.004

Henley, B. J., Thyer, M. A., Kuczera, G., & Franks, S. W. (2011). Climate-informed stochastic hydrological modeling: Incorporating decadal-scale variability using paleo data. *Water Resources Research*, *47*(11), W11509. https://doi.org/10.1029/2010WR010034

Herman, J. D., Quinn, J. D., Steinschneider, S., Giuliani, M., & Fletcher, S. (2020). Climate Adaptation as a Control Problem: Review and Perspectives on Dynamic Water Resources Planning Under Uncertainty. *Water Resources Research*, *56*(2). https://doi.org/10.1029/2019WR025502

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

Hinkley, D. (1977). On Quick Choice of Power Transformation. *Applied Statistics*, *26*(1), 67. https://doi.org/10.2307/2346869

Ho, M., Kiem, A. S., & Verdon-Kidd, D. C. (2015). A paleoclimate rainfall reconstruction in the Murray-Darling Basin (MDB), Australia: 2. Assessing hydroclimatic risk using paleoclimate records of wet and dry epochs. *Water Resources Research*, *51*(10), 8380–8396. https://doi.org/10.1002/2015WR017059

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*(7), 5195–5212. https://doi.org/10.1002/2015WR018444

Ho, M., Lall, U., Sun, X., & Cook, E. R. (2017). Multiscale temporal variability and regional patterns in 555 years of conterminous U.S. streamflow. *Water Resources Research*, *53*(4), 3047–3066. https://doi.org/10.1002/2016WR019632

Holland, J. H. (1975). *Adaptation In Natural and Artificial Systems*. The University of Michigan Press.

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S., & Zhang, H.-M. (2017). Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations,

and Intercomparisons. *Journal of Climate*, *30*(20), 8179–8205. https://doi.org/10.1175/JCLI-D-16-0836.1

Johnson, F., White, C. J., Van Dijk, A. I. J. M., Ekstrom, M., Evans, J. P., Jakob, D., Kiem, A. S., Leonard, M., Rouillard, A., & Westra, S. (2016). Natural hazards in Australia: floods. *Climatic Change*, *139*(1), 21–35. https://doi.org/10.1007/s10584-016-1689-y

Jolliffe, I. (2002). *Principal Component Analysis*. Springer-Verlag. https://doi.org/10.1007/b98835

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

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, *82*(1), 35. https://doi.org/10.1115/1.3662552

Kiem, A. S., Johnson, F., Westra, S., Van Dijk, A. I. J. M., Evans, J. P., O'Donnell, A., Rouillard, A., Barr, C., Tyler, J., Thyer, M. A., Jakob, D., Woldemeskel, F., Sivakumar, B., & Mehrotra, R. (2016). Natural hazards in Australia: droughts. *Climatic Change*, *139*(1), 37–54. https://doi.org/10.1007/s10584-016-1798-7

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323–4331. https://doi.org/10.5194/hess-23-4323-2019

Knoben, W. J. M., Woods, R. A., & Freer, J. E. (2018). A Quantitative Hydrological Climate Classification Evaluated with Independent Streamflow Data. *Water Resources Research*, *54*(7), 5088–5109. https://doi.org/10.1029/2018WR022913

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

Koller, D., & Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.

Komori, D., Nakamura, S., Kiguchi, M., Nishijima, A., Yamazaki, D., Suzuki, S., Kawasaki, A., Oki, K., & Oki, T. (2012). Characteristics of the 2011 Chao Phraya River flood in Central Thailand. *Hydrological Research Letters*, *6*, 41–46. https://doi.org/10.3178/hrl.6.41

Koutsoyiannis, D. (2011). Hurst-Kolmogorov Dynamics and Uncertainty. *Journal of the American Water Resources Association*, *47*(3), 481–495. https://doi.org/10.1111/j.1752-1688.2011.00543.x

Kripalani, R. H., & Kulkarni, A. (1997). Rainfall variability over Southeast Asia—connections with Indian monsoon and ENSO extremes: new perspectives. *International Journal of Climatology*, *17*(11), 1155–1168. https://doi.org/10.1002/(SICI)1097-0088(199709)17:11<1155::AID-JOC188>3.0.CO;2-B

Krishna Kumar, K., Rajagopalan, B., & Cane, M. A. (1999). On the Weakening Relationship Between the Indian Monsoon and ENSO. *Science*, *284*(5423), 2156–2159. https://doi.org/10.1126/science.284.5423.2156

Kummu, M., de Moel, H., Ward, P. J., & Varis, O. (2011). How Close Do We Live to Water? A Global Analysis of Population Distance to Freshwater Bodies (M. Perc, Ed.). *PLoS ONE*, *6*(6), e20578. https://doi.org/10.1371/journal.pone.0020578

Lall, U., & Sharma, A. (1996). A nearest neighbor bootstrap for resampling hydrologic-time series. *Water Resources Research*, *32*(3), 679–693. https://doi.org/10.1029/95WR02966

Lara, A., Bahamondez, A., González-Reyes, A., Muñoz, A. A., Cuq, E., & Ruiz-Gómez, C. (2015). Reconstructing streamflow variation of the Baker River from tree-rings in Northern Patagonia since 1765. *Journal of Hydrology*, *529*(P2), 511–523. https://doi.org/10.1016/j.jhydrol.2014.12.007

Lavigne, F., Degeai, J. P., Komorowski, J. C., Guillet, S., Robert, V., Lahitte, P., Oppenheimer, C., Stoffel, M., Vidal, C. M., Suronoh, Pratomo, I., Wassmer, P., Hajdas, I., Hadmoko, D. S., & De Belizal, E. (2013). Source of the great A.D. 1257 mystery eruption unveiled, Samalas volcano, Rinjani Volcanic Complex, Indonesia. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(42), 16742–16747. https://doi.org/10.1073/pnas.1307520110

Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrological Processes*, *27*(15), 2171–2186. https://doi.org/10.1002/hyp.9740

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., & Wisser, D. (2011). High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment*, *9*(9), 494–502. https://doi.org/10.1890/100125

Li, J., Shao, X., Qin, N., & Li, Y. (2018). Runoff variations at the source of the Yangtze River over the past 639 years based on tree-ring data. *Climate Research*, *75*(2), 131–142. https://doi.org/10.3354/cr01510

Li, J., Xie, S.-P., Cook, E. R., Chen, F., Shi, J., Zhang, D. D., Fang, K., Gou, X., Li, T., Peng, J., Shi, S., & Zhao, Y. (2019). Deciphering Human Contributions to Yellow River Flow Reductions and Downstream Drying Using Centuries-Long Tree Ring Records. *Geophysical Research Letters*, *46*(2), 898–905. https://doi.org/10.1029/2018GL081090

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

Libisch-Lehner, C. P., Nguyen, H. T. T., Taormina, R., Nachtnebel, H. P., & Galelli, S. (2019). On the Value of ENSO State for Urban Water Supply System Operators: Opportunities, Trade-Offs, and Challenges. *Water Resources Research*, *55*(4), 2856–2875. https://doi.org/10.1029/2018WR023622

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* (Vol. 46). 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

Littell, J. S., Pederson, G. T., Gray, S. T., Tjoelker, M., Hamlet, A. F., & Woodhouse, C. A. (2016). Reconstructions of Columbia River Streamflow from Tree-Ring

Chronologies in the Pacific Northwest, USA. *Journal of the American Water Resources Association*, *52*(5), 1121–1141. https://doi.org/10.1111/1752-1688.12442

Liu, Y., Sun, J., Song, H., Cai, Q., Bao, G., & Li, X. (2010). Tree-ring hydrologic reconstructions for the Heihe River watershed, western China since AD 1430. *Water Research*, *44*(9), 2781–2792. https://doi.org/10.1016/J.WATRES.2010.02.013

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

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

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

Lustig, T., & Pottier, C. (2007). *The Angkorian Hydraulic City: Exploitation or Over-Exploitation of the Soil? Translation into English of Groslier , B-P . Le cité hydraulique angkorienne: exploitation ou surexploitation du sol?* (F. Lagirarde, Ed.; Aséanie 20). EFEO Diffusion, Aséanie Bangkok.

Mantua, N. J., & Hare, S. R. (2002). The Pacific Decadal Oscillation. *Journal of Oceanography*, *58*(1), 35–44. https://doi.org/10.1023/A:1015820616384

Marco, J. B., Harboe, R., & Salas, J. D. (Eds.). (1993). *Stochastic Hydrology and its Use in Water Resources Systems Simulation and Optimization*. Springer Netherlands. https://doi.org/10.1007/978-94-011-1697-8

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*(5), 1–12. https://doi.org/10.1029/2010WR010019

Maxwell, R., Harley, G., Maxwell, J., Rayback, S., Pederson, N., Cook, E. R., Barclay, D., Li, W., & Rayburn, J. (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

McPhaden, M. J., Zebiak, S. E., & Glantz, M. H. (2006). ENSO as an Integrating Concept in Earth Science. *Science*, *314*(5806), 1740–1745. https://doi.org/10.1126/science.1132588

Meko, D. M., Touchan, R., Kherchouche, D., & Slimani, S. (2020). Direct Versus Indirect Tree Ring Reconstruction of Annual Discharge of Chemora River, Algeria. *Forests*, *11*(9), 986. https://doi.org/10.3390/f11090986

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

Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, *391*(1-2), 202–216. https://doi.org/10.1016/j.jhydrol.2010.07.012

Morales, M. S., Cook, E. R., Barichivich, J., Christie, D. A., Villalba, R., LeQuesne, C., Srur, A. M., Ferrero, M. E., González-Reyes, Á., Couvreux, F., Matskovsky, V., Aravena, J. C., Lara, A., Mundo, I. A., Rojas, F., Prieto, M. R., Smerdon, J. E., Bianchi, L. O., Masiokas, M. H., … Boninsegna, J. A. (2020). Six hundred years of South American tree rings reveal an increase in severe hydroclimatic events since mid-20th century. *Proceedings of the National Academy of Sciences*, 202002411. https://doi.org/10.1073/pnas.2002411117

Mosteller, F., & Tukey, J. W. (1977). *Data Analysis and Regression: a second course in statistics*. Addison-Wesley.

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

Ng, J. Y., Turner, S. W. D., & Galelli, S. (2017). Influence of El Niño Southern Oscillation on global hydropower production. *Environmental Research Letters*, *12*(3), 034010. https://doi.org/10.1088/1748-9326/aa5ef8

Nguyen, H. T. T. (2020). ldsr: Linear Dynamical System Reconstruction.

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., Galelli, S., Xu, C., & Buckley, B. M. (2020a). Multi-Proxy, Multi-Season Streamflow Reconstruction with Mass Balance Adjustment. *Earth and Space Science Archive*. https://doi.org/10.1002/essoar.10504791.1

Nguyen, H. T. T., Turner, S. W. D., Buckley, B. M., & Galelli, S. (2020b). Coherent stream flow variability in Monsoon Asia over the past eight centuries — links to oceanic drivers. *Water Resources Research*, *e2020WR027*. https://doi.org/10.1029/2020WR027883

Palmer, J. G., Cook, E. R., Turney, C. S. M., Allen, K. J., Fenwick, P., Cook, B. I., O'Donnell, A., Lough, J., Grierson, P., & Baker, P. (2015). Drought variability in the eastern Australia and New Zealand summer drought atlas (ANZDA, CE 1500–2012) modulated by the Interdecadal Pacific Oscillation. *Environmental Research Letters*, *10*(12), 124002. https://doi.org/10.1088/1748-9326/10/12/124002

Palmer, W. C. (1965). *Meteorological Drought. Research Paper No. 45*. U.S. Department of Commerce Weather Bureau.

Panyushkina, I. P., Meko, D. M., Macklin, M. G., Toonen, W. H. J., Mukhamadiev, N. S., Konovalov, V. G., Ashikbaev, N. Z., & Sagitov, A. O. (2018). Runoff variations in Lake Balkhash Basin, Central Asia, 1779–2015, inferred from tree rings. *Climate Dynamics*, *51*(7-8), 3161–3177. https://doi.org/10.1007/s00382-018-4072-z

Patskoski, J., & Sankarasubramanian, A. (2015). Improved reservoir sizing utilizing observed and reconstructed streamflows within a Bayesian combination framework. *Water Resources Research*, *51*(7), 5677–5697. https://doi.org/10.1002/2014WR016189

Pederson, N., Leland, C., Nachin, B., Hessl, A. E., Bell, A. R., Martin-Benito, D., Saladyga, T., Suran, B., Brown, P. M., & Davi, N. K. (2013). Three centuries of shifting hydroclimatic regimes across the Mongolian Breadbasket. *Agricultural and Forest Meteorology*, *178-179*, 10–20. https://doi.org/10.1016/j.agrformet.2012.07.003

Pederson, N., Hessl, A. E., Baatarbileg, N., Anchukaitis, K. J., & Di Cosmo, N. (2014). Pluvials, droughts, the Mongol Empire, and modern Mongolia. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(12), 4375–4379. https://doi.org/10.1073/pnas.1318677111

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, *11*(5), 1633–1644. https://doi.org/10.5194/hess-11-1633-2007

Pelletier, J. D., & Turcotte, D. L. (1997). Long-range persistence in climatological and hydrological time series: Analysis, modeling and application to drought hazard assessment. *Journal of Hydrology*, *203*(1-4), 198–208. https://doi.org/10.1016/S0022-1694(97)00102-9

Pielke, R. A., Wilby, R., Niyogi, D., Hossain, F., Dairuku, K., Adegoke, J., Kallos, G., Seastedt, T., & Suding, K. (2012). Dealing With Complexity and Extreme Events Using a Bottom-Up, Resource-Based Vulnerability Perspective. *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

Ramos, J., Mallants, D., & Feyen, J. (1995). State Space Identification of Linear Deterministic Rainfall-Runoff Models. *Water Resources Research*, *31*(6), 1519–1531. https://doi.org/10.1029/95WR00234

Rao, M. P. (2020). *Hydroclimate variability and environmental change in Eurasia over the past millennium and its impacts* (PhD Thesis). Columbia University. https://doi.org/10.7916/d8-890y-wb66

Rao, M. P., Cook, E. R., Cook, B. I., Palmer, J. G., Uriarte, M., Devineni, N., Lall, U., D'Arrigo, R. D., Woodhouse, C. A., Ahmed, M., Zafar, M. U., Khan, N., Khan, A., & Wahab, M. (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

Räsänen, T. A., & Kummu, M. (2013). Spatiotemporal influences of ENSO on precipitation and flood pulse in the Mekong River Basin. *Journal of Hydrology*, *476*, 154–168. https://doi.org/10.1016/j.jhydrol.2012.10.028

Räsänen, T. A., Lindgren, V., Guillaume, J. H., Buckley, B. M., & Kummu, M. (2016). On the spatial and temporal variability of ENSO precipitation and drought teleconnection in mainland Southeast Asia. *Climate of the Past*, *12*(9), 1889–1905. https://doi.org/10.5194/cp-12-1889-2016

Rauch, H. E., Tung, F., & Striebel, C. T. (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, *3*(8), 1445–1450. https://doi.org/10.2514/3.3166

Razavi, S., Elshorbagy, A., Wheater, H., & Sauchyn, D. (2015). Toward understanding nonstationarity in climate and hydrology through tree ring proxy records. *Water Resources Research*, *51*(3), 1813–1830. https://doi.org/10.1002/2014WR015696

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

Roesch, A., & Schmidbauer, H. (2014). WaveletComp: Computational Wavelet Analysis.

Roweis, S., & Ghahramani, Z. (2001). *Learning Nonlinear Dynamical Systems Using the Expectation-Maximization Algorithm* (Vol. 5). https://doi.org/10.1002/0471221546.ch6

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., & Becker, E. (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

Saito, L., Biondi, F., Devkota, R., Vittori, J., & Salas, J. D. (2015). A water balance approach for reconstructing streamflow using tree-ring proxy records. *Journal of Hydrology*, *529*(P2), 535–547. https://doi.org/10.1016/j.jhydrol.2014.11.022

Saji, N. H., Goswami, B. N., Vinayachandran, P. N., & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, *401*(6751), 360–363. https://doi.org/10.1038/43854

Sano, M., Buckley, B. M., & Sweda, T. (2009). Tree-ring based hydroclimate reconstruction over northern Vietnam from Fokienia hodginsi: Eighteenth century megadrought 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 δ 18 O. *Journal of Geophysical Research: Atmospheres*, *117*(D12), D12115. https://doi.org/10.1029/2012JD017749

Satoh, Y., Kahil, T., Byers, E., Burek, P., Fischer, G., Tramberend, S., Greve, P., Flörke, M., Eisner, S., Hanasaki, N., Magnuszewski, P., Nava, L. F., Cosgrove, W., Langan, S., & Wada, Y. (2017). Multi-model and multi-scenario assessments of Asian water futures: The Water Futures and Solutions (WFaS) initiative. *Earth's Future*, *5*(7), 823–852. https://doi.org/10.1002/2016EF000503

Sauchyn, D., & Ilich, N. (2017). Nine Hundred Years of Weekly Streamflows: Stochastic Downscaling of Ensemble Tree-Ring Reconstructions. *Water Resources Research*, 1–18. https://doi.org/10.1002/2017WR021585

Sauchyn, D., Vanstone, J., St. Jacques, J.-M., & Sauchyn, R. (2015). Dendrohydrology in Canada's western interior and applications to water resource management. *Journal of Hydrology*, *529*, 548–558. https://doi.org/10.1016/j.jhydrol.2014.11.049

Schook, D. M., Friedman, J. M., & Rathburn, S. L. (2016). Flow reconstructions in the Upper Missouri River Basin using riparian tree rings. *Water Resources Research*, *52*(10), 8159–8173. https://doi.org/10.1002/2016WR018845

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).

Schulman, E. (1951). Tree-Ring Indices of Rainfall, Temperature, and River Flow. In T. F. Malone (Ed.), *Compendium of meteorology: Prepared under the direction of the committee on the compendium of meteorology* (pp. 1024–1029). American Meteorological Society. https://doi.org/10.1007/978-1-940033-70-9_82

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

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

Shakun, J. D., & Shaman, J. (2009). Tropical origins of North and South Pacific decadal variability. *Geophysical Research Letters*, *36*(19), L19711. https://doi.org/10.1029/2009GL040313

Shumway, R. H., & Stoffer, D. S. (1982). An Approach to The Time Series Smoothing and Forecasting Using the EM Algorithm. *Journal of Time Series Analysis*, *3*(4), 253–264. https://doi.org/10.1111/j.1467-9892.1982.tb00349.x

Shumway, R. H., & Stoffer, D. S. (2011). *Time Series Analysis and Its Applications With R Examples*. https://doi.org/10.1007/978-1-4419-7865-3

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., Büntgen, U., Caffee, M., Chellman, N., Dahl-Jensen, D., Fischer, H., Kipfstuhl, S., Kostick, C., Maselli, O. J., Mekhaldi, F., Mulvaney, R., Muscheler, R., Pasteris, D. R., Pilcher, J. R., … Woodruff, T. E. (2015). Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature*, *523*(7562), 543–549. https://doi.org/10.1038/nature14565

Singhrattna, N., Rajagopalan, B., Krishna Kumar, K., & Clark, M. P. (2005). Interannual and interdecadal variability of Thailand summer monsoon season. *Journal of Climate*, *18*(11), 1697–1708. https://doi.org/10.1175/JCLI3364.1

Sinha, A., Berkelhammer, M., Stott, L. D., Mudelsee, M., Cheng, H., & Biswas, J. (2011). The leading mode of Indian Summer Monsoon precipitation variability during the last millennium. *Geophysical Research Letters*, *38*(15). https://doi.org/10.1029/2011GL047713

Sinha, A., Cannariato, K. G., Stott, L. D., Cheng, H., Edwards, R. L., Yadava, M. G., Ramesh, R., & Singh, I. B. (2007). A 900-year (600 to 1500 A.D.) record of the Indian summer monsoon precipitation from the core monsoon zone of India. *Geophysical Research Letters*, *34*(16). https://doi.org/10.1029/2007GL030431

Southern Nevada Water Authority. (2012). Colorado River Law.

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., Cook, E. R., Burnette, D. J., Villanueva, J., Cerano, J., Burns, J. N., Griffin, D., Cook, B. I., Acuña, R., Torbenson, M. C., Szejner, P., & Howard, I. M. (2016). The Mexican Drought Atlas: Tree-ring reconstructions of the soil moisture balance during the late pre-Hispanic, colonial, and modern eras. *Quaternary Science Reviews*, *149*, 34–60. https://doi.org/10.1016/j.quascirev.2016.06.018

Stein, L., Pianosi, F., & Woods, R. (2020). Event-based classification for global study of river flood generating processes. *Hydrological Processes*, *34*(7), 1514–1529. https://doi.org/10.1002/hyp.13678

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*.

Stothers, R. B. (1984). The Great Tambora Eruption in 1815 and Its Aftermath. *Science*, *224*(4654), 1191–1198. https://doi.org/10.1126/science.224.4654.1191

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

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

Thyer, M. A., Frost, A. J., & Kuczera, G. (2006). Parameter estimation and model identification for stochastic models of annual hydrological data: Is the observed record long enough? *Journal of Hydrology*, *330*(1-2), 313–328. https://doi.org/10.1016/j.jhydrol.2006.03.029

Tierney, J. E., Abram, N. J., Anchukaitis, K. J., Evans, M. N., Giry, C., Kilbourne, K. H., Saenger, C. P., Wu, H. C., & Zinke, J. (2015). Tropical sea surface temperatures for the past four centuries reconstructed from coral archives. *Paleoceanography*, *30*(3), 226–252. https://doi.org/10.1002/2014PA002717

Tozer, C. R., Kiem, A. S., Vance, T. R., Roberts, J. L., Curran, M. A., & Moy, A. D. (2018). Reconstructing pre-instrumental streamflow in Eastern Australia using a water balance approach. *Journal of Hydrology*, *558*, 632–646. https://doi.org/S0022169418300738

Tozer, C. R., Vance, T. R., Roberts, J. L., Kiem, A. S., Curran, M. A., & Moy, A. D. (2016). An ice core derived 1013-year catchment-scale annual rainfall reconstruction in subtropical eastern Australia. *Hydrology and Earth System Sciences*, *20*(5), 1703–1717. https://doi.org/10.5194/hess-20-1703-2016

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

Tsay, R. S. (2013). *Multivariate Time Series Analysis with R and Financial Applications*. Wiley.

Tsay, R. S., & Wood, D. (2018). *MTS: All-Purpose Toolkit for Analyzing Multivariate Time Series (MTS) and Estimating Multivariate Volatility Models*.

Turner, S. W. D., & Galelli, S. (2016). Regime-shifting streamflow processes: Implications for water supply reservoir operations. *Water Resources Research*, *52*(5), 3984–4002. https://doi.org/10.1002/2015WR017913

Ummenhofer, C. C., Biastoch, A., & Böning, C. W. (2017). Multidecadal indian ocean variability linked to the pacific and implications for preconditioning indian ocean dipole events. *Journal of Climate*, *30*(5), 1739–1751. https://doi.org/10.1175/JCLI-D-16-0200.1

Ummenhofer, C. C., D'Arrigo, R. D., Anchukaitis, K. J., Buckley, B. M., & Cook, E. R. (2013). Links between Indo-Pacific climate variability and drought in the Monsoon Asia Drought Atlas. *Climate Dynamics*, *40*(5-6), 1319–1334. https://doi.org/10.1007/s00382-012-1458-1

van der Schrier, G., Barichivich, J., Briffa, K. R., & Jones, P. D. (2013). A scPDSI-based global data set of dry and wet spells for 1901-2009. *Journal of Geophysical Research: Atmospheres*, *118*(10), 4025–4048. https://doi.org/10.1002/jgrd.50355

Wang, J. K., Johnson, K. R., Borsato, A., Amaya, D. J., Griffiths, M. L., Henderson, G. M., Frisia, S., & Mason, A. (2019). Hydroclimatic variability in Southeast Asia over the past two millennia. *Earth and Planetary Science Letters*, *525*, 115737. https://doi.org/10.1016/j.epsl.2019.115737

Wang, Y., Byers, E., Parkinson, S., Wanders, N., Wada, Y., Mao, J., & Bielicki, J. M. (2019). Vulnerability of existing and planned coal-fired power plants in Developing Asia to changes in climate and water resources. *Energy & Environmental Science*, *12*(10), 3164–3181. https://doi.org/10.1039/C9EE02058F

Watson, E., & Luckman, B. H. (2005). Spatial patterns of preinstrumental moisture variability in the southern Canadian Cordillera. *Journal of Climate*, *18*(15), 2847–2863. https://doi.org/10.1175/JCLI3416.1

Wells, N., Goddard, S., & Hayes, M. J. (2004). A Self-Calibrating Palmer Drought Severity Index. *Journal of Climate*, *17*(12), 2335–2351. https://doi.org/10.1175/1520-0442(2004)017<2335:ASPDSI>2.0.CO;2

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

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

Wu, C. F. J. (1983). On the convergence properties of the EM algorithm. *Annals of Statistics*, *11*(1), 95–103. https://doi.org/10.1214/aos/1176346060

Xu, C., Buckley, B. M., Promchote, P., Wang, S. Y. S., Pumijumnong, N., An, W., Sano, M., Nakatsuka, T., & Guo, Z. (2019). Increased Variability of Thailand's Chao Phraya River Peak Season Flow and Its Association With ENSO Variability: Evidence From Tree Ring δ18O. *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 δ18O-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 δ18O of Fokienia hodginsii in northern Laos: A promising proxy to reconstruct ENSO? *Journal of Geophysical Research: Atmospheres*, *116*(D24), D24109. 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 δ18O. *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., Sun, W., An, W., Wang, W., Zeng, X., & Qin, D. (2019). Regional drought shifts (1710–2010) in East Central Asia and linkages with atmospheric circulation recorded in tree-ring δ18O. *Climate Dynamics*, *52*, 713–727. https://doi.org/10.1007/s00382-018-4215-2

Yang, B., Qin, C., Shi, F., & Sonechkin, D. M. (2012). Tree ring-based annual streamflow reconstruction for the Heihe River in arid northwestern China from ad 575 and its implications for water resource management. *Holocene*, *22*(7), 773–784. https://doi.org/10.1177/0959683611430411

Yang, B., Chen, X., He, Y., Wang, J., & Lai, C. (2019). Reconstruction of annual runoff since CE 1557 using tree-ring chronologies in the upper Lancang-Mekong River basin. *Journal of Hydrology*, *569*, 771–781. https://doi.org/10.1016/j.jhydrol.2018.12.034

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

Yu, E., King, M. P., Sobolowski, S., Otterå, O. H., & Gao, Y. (2018). Asian droughts in the last millennium: a search for robust impacts of Pacific Ocean surface temperature variabilities. *Climate Dynamics*, *50*(11-12), 4671–4689. https://doi.org/10.1007/s00382-017-3897-1

Yuan, Y., Shao, X., Wei, W., Yu, S., Gong, Y., & Trouet, V. (2007). The Potential to Reconstruct Manasi River Streamflow in the Northern Tien Shan Mountains (NW China). *Tree-Ring Research*, *63*(2), 81–93. https://doi.org/10.3959/1536-1098-63.2.81

Zang, C. S. (2015). Dendrobox - An interactive exploration tool for the international tree ring data bank. *Dendrochronologia*, *33*, 31–33. https://doi.org/10.1016/j.dendro.2014.10.002

Zhang, R., Qin, L., Yuan, Y., Gou, X., Zou, C., Yang, Q., Shang, H., & Fan, Z. (2016a). Radial growth response of Populus xjrtyschensis to environmental factors and a century-long reconstruction of summer streamflow for the Tuoshigan River, northwestern China. *Ecological Indicators*, *71*, 191–197. https://doi.org/10.1016/j.ecolind.2016.06.035

Zhang, R., Yuan, Y., Gou, X., Yang, Q., Wei, W., Yu, S., Zhang, T., Shang, H., Chen, F., Fan, Z., & Qin, L. (2016b). Streamflow variability for the Aksu River on the southern slopes of the Tien Shan inferred from tree ring records. *Quaternary Research*, *85*(3), 371–379. https://doi.org/10.1016/j.yqres.2016.03.001

Zhang, T., Yuan, Y., Chen, F., Yu, S., Zhang, R., Qin, L., & Jiang, S. (2018). Reconstruction of hydrological changes based on tree-ring data of the Haba River, northwestern China. *Journal of Arid Land*, *10*(1), 53–67. https://doi.org/10.1007/s40333-017-0034-2

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), D11307. https://doi.org/10.1029/2011JD017198