Machine Learning-enhanced Hydrologic Modeling under Changing Climate

by

Qianqiu Longyang

A Dissertation Presented in Partial Fulfillment
of the Requirement for the Degree
Doctor of Philosophy

Approved March 2024 by the
Graduate Supervisory Committee:

Ruijie Zeng, Chair
Giuseppe Mascaro
Sebastien Motsch
Tianfang Xu
Zhihua Wang

ARIZONA STATE UNIVERSITY

May 2024

ABSTRACT

Hydrological modeling has been widely used to predict the response of a hydrologic system to changing drivers in short to long terms, thus providing quantitative decision-making support for strategic planning and adaptation policies development. Despite advancements in hydrological modeling, challenges remain for improving the quality and realism of hydrologic predictions. Machine Learning (ML), known for its proficiency in extracting patterns from large datasets, has attracted interest within the hydrology community and has been applied to tackle various challenges in predicting the hydrologic cycle over the past few years.

This dissertation focuses on the enhancement of hydrologic models through ML and evolving datasets, with the goal of improving their accuracy and utility. The first part addresses the issue of missing or inadequately represented components in the existing large-scale hydrologic models. Taking the simulation of regulated flow conditions as an example, the study presents a hierarchical temporal scale framework for all data-driven reservoir release models, enabling more effective use of limited data sources and ensuring that practical significance aligns with model configuration. The second part addresses the computational challenges associated with model parameterization. The development of an ML-based surrogate model expedites the parameter estimation and calibration, particularly for those properties that may vary over time and require the adoption of dynamic parameterization. Satellite-derived vegetation interannual variability serves as a case study in this dissertation, illustrating how the dynamic nature of vegetation can influence hydrologic responses. From the perspective of hydrologic modelers, these two parts of work enhance the hydrologic model's realism by improving both its representation and parameterization, respectively. For water managers, a combination of surrogate model and the reservoir operation module enables integrated reservoir management modeling under different climate projection scenarios. Building upon the insights gained from the first two

parts, the last part shows such application to translate hydrologic and climatic data into actionable strategies for water management. Utilizing 21 federal reservoirs in Texas as a case study, this part offers a framework for stakeholders to assess the effectiveness of current reservoir operation policies under future climate scenarios through the interactions among hydroclimatology, reservoir infrastructure, and operation policy.

ACKNOWLEDGEMENTS

First and foremost, I extend my deepest and most heartfelt gratitude to my advisor and dissertation committee chair, Prof. Ruijie Zeng. Prof. Zeng is nothing short of a genius, full of brilliant ideas and passion. His expertise spans across various research domains, always with an eye on the broader vision while maintaining an uncompromising attention to detail. From our very first meeting on January 12, 2019, his invaluable guidance, unwavering support, and warm encouragement have been the cornerstone of my journey. He has been instrumental in my evolution from a geology undergraduate to a Ph.D. in engineering.

I am profoundly thankful to my Ph.D. committee members, Prof. Giuseppe Mascaro, Prof. Sebastien Motsch, Prof. Tianfang Xu, and Prof. Zhihua Wang, for their inspiring discussions, robust support, and the invaluable, insightful feedback on my research. Prof. Mascaro, who taught me my first engineering and watershed modeling courses, has always been eager to answer our detailed questions and assist students. Prof. Motsch, who taught me about deep neural networks, has consistently shown endless patience and provided insightful guidance, helping me to grasp the intricacies of deep learning. My sincerest thanks to Prof. Tianfang Xu, who was instrumental particularly during the "training" stages of my Ph.D. Her patient and constructive feedback on the technical details of my work has been immensely beneficial. I am profoundly indebted to Prof. Zhihua Wang, who has offered encouragement and wise advice throughout my doctoral journey. His teachings, which I have been fortunate to absorb throughout my Ph.D., have shaped my academic pursuits and personal growth. The contributions of each committee member have been pivotal to my development, and for this, I am eternally grateful.

I also seize this moment to express my appreciation to all the faculty at the School of Sustainable Engineering and the Built Environment for their help and advice. Additionally, I want to take this opportunity to wholeheartedly thank all my lab mates, especially Dr.

Zhaocheng Wang, Dr. Haowen Yue, Dr. Xueli Yang, Dr. Xin Su, Dr. Mu Xiao, Dr. Haoran Hou, and the future Drs. Shiqi Wei, Chuncheng Yao, Qinyuan Dai, Yihang Wang. Their kindness and companionship (especially during our memorable hot pot gatherings) have not only made my Ph.D. journey more manageable but profoundly enjoyable as well.

A special word of thanks goes to my parents, whose endless love has been my unwavering source of strength. Lastly, I am grateful to my best friends, the future Drs. Maorui Cai, Jiao Bai, and Xu Liu, for their support throughout this remarkable journey.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Chapter 1

INTRODUCTION

1.1 Background and Motivation

Hydrologic models and data serve as the backbone for sustainable water management (Jakeman *et al.*, 2006; McMillan *et al.*, 2018; Ward *et al.*, 2019; Hadjimichael *et al.*, 2023). Hydrologic models, by quantifying hydrologic fluxes (e.g., precipitation, evapotranspiration, interception, infiltration, runoff) and states (e.g., soil moisture, snow water equivalent), enable water availability estimates, risk assessment of floods and droughts, and strategic planning and adaptation policies development (Brown *et al.*, 2015; Samaniego *et al.*, 2019). Hydrologic data provide observations into hydrologic processes, ensuring models' fidelity to real-world scenarios. Hydrologic models and data are fundamentally interconnected. Models are grounded in physical laws that are themselves intrinsically derived from observations, and evaluated and constrained against observations; concurrently, models function as interpolators, adeptly bridging gaps in our observations, which are typically limited to specific variables and frequently sparse both spatially and temporally (Gettelman *et al.*, 2022). In a changing future where the assumption of stationarity in climate and hydrological systems may become untenable, models require ongoing validation and updating, synthesizing new observations to maintain their realism and adapt to these changes (Milly *et al.*, 2008b; Van Dijk, 2011). Therefore, the seamless integration of hydrologic models and data is not just essential for understanding the hydrologic system but also critical for adapting modeling approaches to the evolving climate, thereby facilitating informed and responsive decision-making.

Surging models and data may overwhelm hydrologic forecasting and prediction systems.

Over the past few decades, substantial advancements in computational power, along with a deeper understanding of the underlying mechanism of hydrologic systems, have propelled the development of increasingly sophisticated hydrologic models (Liu and Gupta, 2007). The advent of advanced satellite, airborne, and ground-based remote sensing has led to a dramatic surge in hydrologic data repositories (Butler, 2014; Chen and Wang, 2018). This proliferation of models and data, while marking significant progress, brings forth a unique set of challenges for improving the quality and realism of hydrologic predictions.

A central challenge lies in the paradox of abundance: on the one hand, there exists a surplus of models that, despite their quantity, consistently fall short of accurately capturing some key responses, indicating insufficient scientific understanding of hydrologic dynamics (Clark *et al.*, 2011). For instance, some processes involved with anthropogenic factors that alter the hydrological regimes, are missing or inadequately represented in the current hydrologic modeling. This shortfall is not due to a lack of models but stems from fundamental limitations, such as the scarcity of observations and the complexity of quantitatively characterizing human impacts on the hydrology system, which may hinder the quality of seasonal forecasts and the usability of the climate projections (Haddeland *et al.*, 2006; Samaniego *et al.*, 2019; Singh and Basu, 2022). On the other hand, there is an overwhelming accumulation of data that, ironically, has not been effectively assimilated and comprehended to enhance these models (Reichstein *et al.*, 2019). The challenge highlights the need for leveraging a large amount of data to fill critical gaps in existing models, especially in the representation of human-water interactions.

The second major challenge is the computationally demanding process of hydrologic model parameterization, which is further complicated by the intricacy of sophisticated models and the vastness of the datasets involved. Process-based distributed hydrologic models put forward high requirements for parameterization to serve accurate predictions, commonly involving resource-intensive computational efforts (Boyle *et al.*, 2000). Limited

computational resources may impede the process of parameter estimate and calibration. Hydrologic parameters can be categorized into two types: physically significant, measurable parameters (e.g., Leaf Area Index) and empirically derived parameters (e.g., the Clapp and Hornberger "B" soil parameter to determine hydraulic conductivity) (Clapp and Hornberger, 1978). The parameters with physical meaning could be estimated using remote sensing datasets. Nevertheless, the availability of various remotely sensed datasets means that integrating satellite-derived spatiotemporal parameter information into hydrologic models, as well as comparing and selecting the optimal products, becomes more time-consuming and labor-intensive. Parameters that are either physically meaningful but challenging to measure (e.g., many subsurface properties), or those that are empirically derived, are typically determined through parameter calibration and regionalization. The calibration process involves adjusting the model parameters at each location to minimize the discrepancies between model outputs and observational data, which requires thousands of model iterations for calibrating a relatively small set of parameters (Tsai *et al.*, 2021). Additionally, the necessity for model recalibration upon updating parameter information introduces further complexity, often constrained by high computational expenses. To overcome these challenges, a tool that accelerates model parameterization is required, one that strikes a balance between accuracy and computational efficiency.

Another key ongoing challenge arises from quantifying the potential impacts of climate change on existing water infrastructure and operational strategies and devising effective risk mitigation approaches. There is a pressing need for managers to continually adapt and refine their understanding in response to evolving climate projections. For instance, the Coupled Model Intercomparison Project, a framework for upgrading knowledge of climate change, has undergone significant developments, from Phase 3 (CMIP3) starting in 2000, Phase 5 (CMIP5) in 2006, to Phase 6 (CMIP6) in 2015. Notably, CMIP3 was marked by a higher underestimation of warming trends, while CMIP5 is more in line

with the observations (Carvalho *et al.*, 2022), and CMIP6 has offered projections with a narrower uncertainty margin over Northern America (Martel *et al.*, 2022). This progression towards more realistic climate projections underscores the critical need for water managers and stakeholders to continually refresh and update their perspectives on climate dynamics. Crucially, these updates are expected to be integrated into hydrologic modeling, along with advancements in land surface processes understanding (e.g., more accurate representations of human regulations and vegetation dynamics). Such integration can enable effectively translating broad-scale atmospheric processes, typically modeled at resolutions of hundreds of kilometers, into detailed, node-based analyses at a finer scale, ensuring that water management strategies remain effective and responsive to the latest scientific understanding.

Recent advances in statistical modeling and Machine Learning (ML) present exciting new opportunities for enhancing hydrologic modeling in the context of climate change (Reichstein *et al.*, 2019). ML, for example, excels at extracting quantitative patterns or modes from large datasets, effectively supplementing model components that have highly uncertain process formulations and are challenging to represent accurately with a physical process-based perspective (Xu and Liang, 2021). A notable instance of this is the two-way feedback between human and water systems (Meempatta *et al.*, 2019). Additionally, ML leverages advancements in computing technology (e.g., Graphical Processing Units) to facilitate rapid diagnosis and accelerate model calibration, which could substantially reduce the computational load involved in parameter estimation and uncertainty quantification (Xu and Valocchi, 2015; Xu *et al.*, 2017). Building on these advancements, ML is poised to play a crucial role in assessing the impacts of climate change on existing water management practices and offering actionable insights for decision-making (Ehsani *et al.*, 2017; Secci *et al.*, 2023).

1.2 Objectives

The overarching goal of this dissertation is to **harness the power of ML to enhance the accuracy and utility of hydrologic modeling, thereby facilitating informed decision-making in surface water resource management amidst changing climate conditions**.

To realize this goal, the research is structured around specific tasks:

1) Inferring the multi-scale representations of human regulations on streamflow from the observed reservoir operation data records. This task is designed to enhance large-scale hydrologic models by effectively utilizing information across multiple temporal scales, derived from the available timeseries data.

2) Incorporating satellite-derived vegetation interannual variability into land surface models (LSM). This involves building an ML-based tool aimed at streamlining parameter calibration and facilitating data fusion within LSMs.

3) Building upon the insights gained from the first two tasks, the final task involves devising a flexible framework. This framework enables assessing the impacts of climate change on existing water infrastructure and operational strategies and identifying opportunities for reoperation in the context of a changing climate.

Through these specific tasks, this dissertation will demonstrate the integration of data-derived insights through ML to fill missing modeling components, expedite model parameterization, and convert these findings into practical, actionable strategies in water resource management.

1.3 Organization of the Dissertation

The rest of this dissertation is organized into four main parts, each aligning with the stated objectives and tasks. Chapter 2 introduces a novel framework designed to enhance

the representation of regulated flow conditions by capturing multiple temporal scales of information. It hypothesizes that incorporating this multi-temporal information will improve the simulation of regulated flow conditions. This hypothesis is tested across more than 300 cases within the contiguous United States (CONUS). This chapter was previously published by *Water Resources Research* (Longyang and Zeng, 2023a). In Chapter 3, a ML-based surrogate model is developed to replicate the physics of LSM. This model facilitates parameter calibration and data fusion. By integrating satellite-derived vegetation dynamics that vary annually into the LSM (or its surrogate), the chapter uses the Upper Colorado River Basin (UCRB) as a case study to demonstrate the importance of considering vegetation variability in water and energy budget estimates, especially for assessing climate change impacts on the water cycle. Chapter 4 builds on the findings from Chapters 2 and 3. It proposes a flexible framework that utilizes the calibrated hydrologic model (or its surrogate) and data-driven regulated flow model to investigate the impact of climate change on flow regulations. It also explores modifications in water management strategies. The chapter employs Texas reservoirs as test cases to illustrate the translation of climate information into actionable strategies using ML. Finally, Chapter 5 concludes the dissertation by summarizing the key findings, particularly focusing on the role of ML in enhancing hydrologic modeling. It also discusses the broader implications of this work and offers recommendations for future research directions.

Chapter 2

INFERRING MULTI-SCALE HUMAN REGULATIONS ON STREAMFLOW FROM OBSERVED RESERVOIR OPERATION DATA

2.1 Introduction

Anthropogenic activities, such as reservoir operation (Biemans *et al.*, 2011; Döll *et al.*, 2009; Haddeland *et al.*, 2006; Zeng and Ren, 2022; Zhao *et al.*, 2021), urbanization (Li *et al.*, 2020; Oudin *et al.*, 2018), and large-scale irrigation (Condon and Maxwell, 2019; Ferguson and Maxwell, 2011; Siebert *et al.*, 2010), have become increasingly important or even dominant driving forces of hydrologic processes in many watersheds over the world. In these watersheds, the streamflow observed at gauging stations represents the interaction between hydrologic and anthropogenic driving forces, rather than the "natural" or "unregulated" flows simulated in hydrologic models (Blair and Buytaert, 2016; Clark *et al.*, 2015). Reservoirs are one of the key water infrastructures that directly regulate the streamflow timing and variability to fulfill various purposes including flood control, water supply, hydroelectricity generation, navigation, and fluvial ecosystem services (Boulange *et al.*, 2021; Ehsani *et al.*, 2017; Forsberg *et al.*, 2017; Lehner *et al.*, 2011; Moran *et al.*, 2018; Ortiz-Partida *et al.*, 2016; Patterson and Doyle, 2018; Simonovic, 1992). In the US, the National Inventory of Dams reports that there are more than 90,000 reservoirs (defined as equal or exceed 25 feet in height and exceed 15 acer-feet in storage, or exceed 6 feet in height and equal or exceed 50 acer-feet storage) regulating the streamflow (DeNeale *et al.*, 2019). These reservoirs altogether store freshwater resources equivalent to 1 year's average natural runoff (Graf, 1999), generates about 6.3% of total electricity and 31.3% of renewable energy production (Energy Information Administration, 2022), and protect hundreds of

millions of populations from flooding. Meanwhile, the current reservoir operation policies are challenged by shifting flow conditions under climate change (Boulange *et al.*, 2021), elevated risks due to aging infrastructure (Lane, 2007), increasing demand for water supply reliability, and the need for aquatic habitat restoration (Palmer and Ruhi, 2019; Tonkin *et al.*, 2018). Understanding how reservoirs are operated and their interaction with hydrologic cycle is vitally important for assessing the reliability and risks of reservoir functioning (Brekke *et al.*, 2009), designing adaptation strategies for future climate (Ho *et al.*, 2017), and mitigating the tradeoffs among conflicting operation targets (Chen and Olden, 2017; Giuliani *et al.*, 2021; Suen and Eheart, 2006) to achieve sustainable water resources management.

Reservoirs are decision hubs that integrate the complex feedback between hydrologic variability and operational targets under various constraints, such as reservoir inflow, water storage capacity, hydroelectricity generation requirement, and competition among different operation purposes. Challenges remain in modeling the reservoir release decisions, which often involve complex and undocumented decision processes. Often, reservoir operation guidelines are based on predefined rule curves (Klipsch and Hurst, 2007; Yates *et al.*, 2005), which determine release decisions based on water availability, which in turn, depends on inflow and storage (Chen *et al.*, 2022a,b). However, many reservoirs are actively managed, where the flow releases are determined by reservoir managers to account for the complex tradeoffs among different operation targets. This complicated decision-making process often cannot be described with simple operation rules. In addition, observations on reservoir operation (e.g., reservoir water level and release) are very limited due to the complex ownership and regulations.

As a result, reservoirs, as coupled natural-human systems (Liu *et al.*, 2007), are not adequately represented in current hydrologic or hydraulic models. Compared to natural hydrologic processes that can be expressed by physical relationships, it remains unclear how reservoirs are operated to regulate streamflow, as observations on reservoir operation

(e.g., reservoir water level and release) are very limited due to the complex ownership and regulations. For example, the National Water Model is able to predict streamflow for over 2 million reaches in the US, while a limited number of reservoirs are simulated by a simple level pool routing scheme (Gochis *et al.*, 2018; Khazaei *et al.*, 2021) where reservoir releases are passively determined by reservoir water level and spillway characteristics based on hydraulic laws (e.g., weir flow equations). However, the releases from actively managed reservoirs, which are crucial infrastructures involving multiple stakeholders and with significant downstream impacts, are regulated by gates and determined by reservoir managers based on a range of real-world constraints and trade-offs.

Traditionally, reservoir operation rules have been derived using optimization techniques. These models aim to determine optimal releases to achieve predefined objectives (such as minimizing flood risk or maximizing water supply reliability) under various constraints (such as reservoir storage capacity and allowable downstream release). However, actual reservoir release usually deviates from the optimized prescription due to several limitations. First, the theoretical optimal reservoir releases are obtained under a small set of predefined objectives and constraints, which often do not capture the full spectrum of real-world operation conditions (Giuliani *et al.*, 2021). Second, reservoir characteristics (storage capacity vs. water level relationship) or streamflow regime may be different from the conditions when the optimal operation rule was derived. Third, optimization models assume perfect streamflow predictions or a known streamflow prediction uncertainty, but it is not necessarily the case that streamflow prediction is available for operational purposes and whether reservoir managers utilize the streamflow prediction during the decision-making processes (Zhao *et al.*, 2011). Therefore, with these deviations from assumptions, optimization model-derived reservoir operation rules may provide valuable normative solutions for the large-scale hydrologic and water resource model, but often fail to yield satisfactory results for predicting streamflow downstream of reservoirs.

Data-driven models (DDMs) offer a promising alternative to derive reservoir operation rules from historical records of hydrologic and reservoir data (Aboutalebi *et al.*, 2015; Hipni *et al.*, 2013; Lin *et al.*, 2006; Turner *et al.*, 2020a,b; Wei and Hsu, 2008; Yang *et al.*, 2017; Zhang *et al.*, 2018; Zhao and Cai, 2020). Recent studies have demonstrated the capability of various machine learning techniques in capturing reservoir release decisions (Chen *et al.*, 2022a,b; Coerver *et al.*, 2018; Dong *et al.*, 2023; Gangrade *et al.*, 2022; Mateo *et al.*, 2014; Yassin *et al.*, 2019). The rationale is straightforward: if a manager determines the reservoir releases based on some principles (either empirical or optimal) depending on hydroclimatic variation, DDMs can recover the patterns of operation from the reservoir records and other hydroclimatic variables. In addition, compared to optimization models, DDMs are computationally efficient and readily coupled with hydrologic and hydraulic models. The primary motivation behind this chapter is to contribute to the development of simulation strategies that can enhance the representation of reservoirs in regional or national-scale hydrological models, such as the National Water Model.

This chapter hypothesizes that reservoir operation patterns vary across time scales, thus requiring a hierarchical temporal scale configuration of DDMs. First, reservoirs usually have multiple operation purposes that require decisions made at different time scales. For example, daily or hourly release decisions are made for hydroelectricity generation based on the demand from power grids, while the reservoir storage for agricultural water supply exhibits a slow-varying seasonal pattern. Even for reservoirs with one primary operation purpose, hydroclimatic variabilities at different time scales may lead to different operation decisions. A reservoir designed for flood control may be actively operated only during wet seasons to mitigate floods, and the storage may remain relatively stable during dry seasons. Second, release decisions for different operational purposes are made based on different information that changes with time scales. For example, flood control decisions may depend on current reservoir water level and streamflow forecast with leading time up to

several days, while water supply reservoirs may ignore the short-term streamflow variability and focus on hydrologic seasonal dynamics such as snowpack. Third, operation decisions made at different scales interact with each other. The flood control hourly operations during a high flow event may be constrained water level set by seasonal water supply targets; flood control operations, in return, determine initial water level for water supply release for the next decision period. Based on these observations, capturing the reservoir operation decisions across time scales is essential to accurately represent the anthropogenic regulation on streamflow variability.

Despite significant progress in data-driven reservoir modeling, current approaches typically rely on a single time scale for operations, with limited exploration of frameworks that account for multi-timescale interactions. For instance, Zhang *et al.* (2018) assessed the performances of various DDMs with different time resolutions (e.g., hourly, daily, and monthly) for Gezhouba Dam, while neglecting the interactions of decision-making processes across time scales. Yang *et al.* (2021) provided a comprehensive comparison of different DDMs to simulate the daily reservoir outflow over the Upper Colorado Region using the daily inflow, storage, and calendar time as model inputs, which did not completely include decision variables at monthly scales. Turner *et al.* (2020b) built a daily scale DDM for reservoirs in the Columbia River basins with seasonally varying relations that specify water release as a function of prevailing storage levels and forecasted future inflow. However, this approach is based on pre-assumed linear piecewise relations to represent the seasonality, which still needs to be specified based on the modeler's assumption. While single-scale models may adequately serve the needs of reservoir operators, investors, and decision makers for simpler reservoir systems, multi-objective reservoirs and multi-reservoir systems demand greater attention to the full range of timescales for improved reservoir operation modeling. The study conducted by Hejazi *et al.* (2008) using weekly/monthly datasets revealed that the importance of hydrologic indicators vary across

seasons and purposes (i.e., flood control, water supply, hydropower, and irrigation) for reservoirs located in California and Great Plains regions. It highlighted the interdependence between decision variables, purposes and time scales in reservoir operations. The time-varying sensitivity analysis at daily scale for a multi-reservoir system in the Red River Basin further illustrated that effective operating policies adapt the utilization of information over time while coordinating it across multiple reservoirs (Quinn *et al.*, 2019). The challenges arise when simulating regulated flow downstream of such complex reservoirs. A general and flexible framework is needed, which can effectively simulate the reservoir release decisions and capture trade-offs among multiple reservoir operation objectives, as well as the interactions between hydroclimatic conditions and human decisions across various time scales. Furthermore, this framework is expected to be readily compatible with large-scale hydrologic and water resource management models.

This chapter develops a hierarchical temporal scale framework to model reservoir operation decisions across various time scales. The proposed framework exhibits generality in several aspects: (a) it does not require prior knowledge of reservoir operation objectives; (b) it supports the implementation of diverse data-driven modeling techniques; and (c) it utilizes commonly available datasets for training the machine learning models. The framework has the flexibility to (a) use time scale-specific inputs for DDMs to learn reservoir operation behaviors pertinent to each time scale and (b) enable decisions at different time scales to interact with each other. This chapter demonstrates the framework with a two-layer configuration, at monthly/weekly and daily scales, respectively. The framework is validated using the daily operational records of 327 major reservoirs in the United States regulated by the United States Army Corps of Engineers (USACE) and the United States Bureau of Reclamation (USBR). These reservoirs cover a wide spectrum of hydroclimatic conditions, reservoir characteristics and operation purposes, therefore can examine the robustness of the proposed hierarchical temporal scale framework. The monthly-/weekly-scale DDM

learns reservoir decisions unaffected by short-term variability and provides constraints for the daily scale model which captures the event-scale operation rule that deviates from the monthly/weekly average. This framework is flexible to incorporate additional temporal layers (such as at hourly or seasonal scales). This chapter further evaluates which variables are dominant for reservoir operations across various time scales and investigate the tradeoff between training variables and modeling temporal resolution in representing reservoir decisions.

2.2 Methods

2.2.1 Hierarchical Temporal Scale Configuration of Data-driven Models (DDMs)

This chapter models reservoir release schemes at each temporal scale (e.g., daily, weekly, monthly) collectively under a set of hydroclimatic explanatory variables (e.g., streamflow, precipitation). The raw daily time series is separated into a coarse time-scale average (i.e., monthly as illustrated in the example in Figure 2.1) and a fine-scale "deviation." The "deviation" $\hat{y}^{(D)}$ between the daily scale release $y^{(D)}$ from the monthly scale release $y^{(M)}$ is defined as

$$\hat{y}^{(D)} = y^{(M)} - y^{(D)} \tag{2.1}$$

The "deviation" $\hat{y}^{(D)}$ includes (a) true signals (systematic bias or structured error) resulting from fine time scale reservoir release deviating from a coarse time scale operation (e.g., operation for daily release for flood control constrained by monthly water storage target for water supply) and (b) unstructured random error (e.g., Gaussian type random noise from measurement error). This chapter hypothesizes that the structured error between different time scales of observed releases contains information that is not adequately represented at a single time scale, which can be effectively modeled using a hierarchical approach. For

example, the author finds that temporal autocorrelation of the deviations of reservoir releases between daily and weekly/monthly scales exists in most of the reservoirs, probably indicating that relying solely on monthly/weekly averages may not fully capture the intricacies of reservoir release dynamics. This chapter utilizes coarse-scale averages as a source of long-term information to compensate for the limited forward-looking capacity of fine-scale limited time steps.

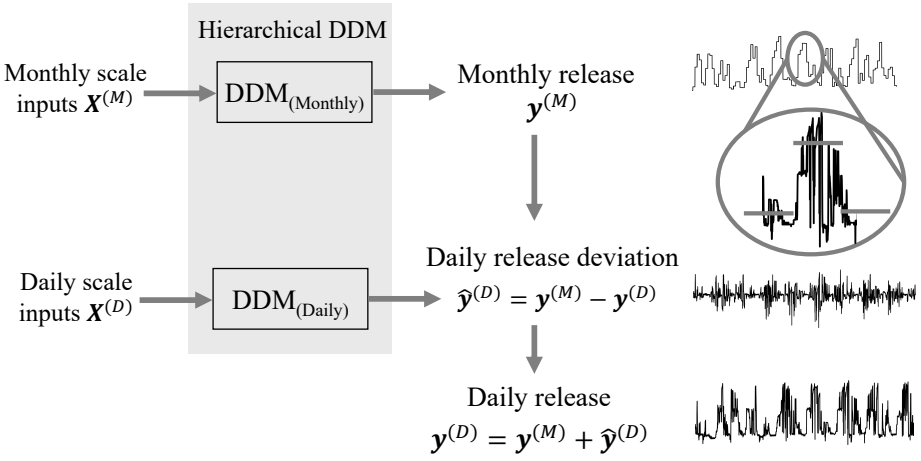

Figure 2.1: The hierarchical temporal scale framework with two layers shown for illustration. The top layer uses a monthly Data-driven model (DDM) to simulate monthly averaged releases ($y^{(M)}$), and the subsequent bottom layer uses a daily DDM to simulate the daily deviation $\hat{y}^{(D)}$, or the difference between daily $y^{(D)}$ and monthly averaged $y^{(M)}$ releases.

The hierarchical temporal scale framework (shown in Figure 2.1) consists of multiple layers, where each layer has a DDM to learn the reservoir operation rules at the corresponding time scale (e.g., monthly, weekly, and daily). The configuration starts from the upper layer corresponding to a coarse time scale (i.e., monthly/weekly in this chapter) to capture the reservoir operation behaviors under slow-varying targets (e.g., storing water for growing season irrigation supply). Historical hydroclimate and reservoir records are aggregated into monthly/weekly time series to train a DDM. The lower layer refines the model to a fine

time scale (i.e., daily scale in this chapter), and a second DDM is trained to simulate the "deviation," defined as the difference between the fine-scale release and release simulated by the coarse time scale DDM. The deviation characterizes short-term deviations from release determined under long-term operation targets and may be caused by gaps between planned and actual situations and complicated tradeoffs between various purposes served in different periods. It is worth noting that the deviation $\hat{y}^{(D)}$ could be defined as the differences between observed releases at a coarse and a fine scale. The difference lies in the fact that the former, which defines the target deviation $\hat{y}^{(D)}$ as the difference between the observed daily release and simulated monthly/weekly release, to some extent, resembles the concept of a boosting algorithm, where the model is improved through the combination of multiple weak models to form a strong model, whereas the latter purely integrates multi-timescale information to generate the target fine-scale release. The effects of the two are considered equivalent when the model in the first layer is able to accurately predict the target release at the coarse scale.

The hierarchical configuration of the framework is flexible to add layers as needed to represent operation decisions at coarser (e.g., seasonal) or finer time scales (e.g., flood control release or hydroelectricity generation under power grid demand) if reservoir operation record is available. In addition, the hierarchical framework allows models at each time scale to take different training variables since different operations decisions may depend on different information. For example, the operation for irrigation water supply may mainly depend on the crop water demand during the growing season, while the operation for flood control may depend on the current reservoir water level and upstream flow predictions for the next few days. By learning the deviations between water release at fine time scale and the coarse time scale average, the DDM can capture the interactions of operation rules at different time scales and represent the tradeoffs between various operation targets. For example, the release for flood control may be dependent on the current reservoir water level, which is affected by the storage target for water supply determined 1 month ago. The reservoir water

level after flood control release may further affect water supply decisions in future time steps. Therefore, the deviation between two layers (i.e., two temporal scales) may represent the tradeoffs between various operation targets.

Two distinct strategies can be employed to train the DDM in each layer: "iterative" and "detached." The iterative strategy enables concurrent updates to all temporal layers throughout the model training process. For neural network-type models such as Multi-Layer Perceptron and Recurrent Neural Networks (RNN), a loss function that spans all temporal scales or multiple loss functions for each temporal scale can be defined, and weight updates are executed in each training epoch. The detached approach involves a simple arithmetic summation or weighted aggregation of the outputs from all layers to generate the final simulations. This chapter uses the iterative strategy to train the DDM.

2.2.2 Hydroclimatic and Reservoir Data

This chapter applies the proposed framework to 248 reservoirs operated by the USACE and 79 reservoirs operated by the USBR across the Contiguous United States (CONUS). These reservoirs are generally actively managed reservoirs with multiple designed purposes. The standardized database for historical daily reservoir levels and operations of USACE reservoirs is developed by (Patterson and Doyle, 2018), while that of USBR reservoirs is accessed via Reclamation Information Sharing Environment. The author sources some data from ResOpsUS, a comprehensive data set on historical reservoir operations in the United States that was recently published by (Steyaert *et al.*, 2022). These observed records include daily reservoir water elevation (feet, ft), storage volume (acer-feet, af), inflow (cubic feet per second, cfs), and release (cubic feet per second, cfs) for each reservoir, with different record lengths and intermittent gaps in the middle of the record due to data collection issues. All reservoirs with continuous records are included in this chapter. For some reservoirs with missing data during only a short period of time (less than 5 days), the nearest neighbor

interpolation method is applied to fill in these gaps to obtain a continuous record. Overall, the continuous records have an average length of 30 years.

The reservoir release data is used as target (response variable) to train and test the DDMs, and water storage volume, reservoir inflow records and hydroclimatic data are used as inputs. The daily-scale meteorological forcing, including total precipitation rate (P, mm/day), potential evapotranspiration (PET, mm/day), and air temperature (T, °C) are obtained from the North American Land Data Assimilation System forcing (Xia *et al.*, 2012). The hydroclimatic data are averagely aggregated over the catchment area upstream of the reservoir to encapsulate the local weather information relevant to reservoir operation. Specifically, the PET represents atmospheric demand for reservoir evaporative loss, which is substantial for reservoirs in arid and semi-arid regions (Friedrich *et al.*, 2018). The P may reflect the local runoff contribution to the reservoir, while the reservoir inflow represents the runoff from the larger upstream contributing area. The difference between P and PET captures the crop irrigation water demand (Le Page *et al.*, 2021), which may provide important information for reservoirs with irrigation water supply purposes. The gridded snow depth (SD, mm) data retrieved from Broxton *et al.* (2019) is aggregated over the catchment area upstream of the reservoir to account for changes in snowmelt contributions over time. Depending on the specifics of a given reservoir, other information (e.g., hydroelectricity generation) can also be fed into DDMs as inputs.

2.2.3 Experimental Setup

Three groups of experiments are carried out to assess the performances of data-driven reservoir operation models with (a) under different time scale configurations and (b) different combinations of input variables (Table 2.1). The experimental setup is summarized in Table 2.1. The first group of experiments simulate reservoir release solely on a single daily scale (i.e., daily inputs are employed to model the daily release). This strategy is commonly

implemented in existing machine learning based reservoir models. The other two groups of experiments adopt a two-level hierarchical time scale configuration. The second group of experiments receives weekly-average input variables in the first layer to generate weekly average release, and then use daily inputs to model the deviation (difference between daily release and weekly average) in the second layer, herein referred to as "Weekly-Daily (WD)." Similarly, the third group of experiments simulates monthly scale reservoir release in the first layer and refines reservoir release on the daily scale in the second layer, referred to as "Monthly-Daily (MD)." On the daily scale, the author uses the 7 days in the past of input variables to determine release on a given day. For the WD and MD models, the coarse-resolution input variables of the past four steps (weeks or months) are used to derive the release at the current time step, and the daily scale deviations are simulated with daily input variables of the past 7 days. While inflow forecasts have been proven to strongly influence seasonal reservoir operations, particularly for the high-elevation reservoirs fed by snowmelt in the western United States Turner *et al.* (2020a), this chapter only uses the observed records in the past time steps, since it is difficult to acquire the actual streamflow forecasts for each reservoir in the historical period.

To explore the importance of each input variable for predicting reservoir operation at various time scales, this chapter develops six experiments by varying the combinations of input variables in the three groups (Table 2.1). In Experiment 1, daily observed reservoir inflow (I), water storage (S), and hydroclimatic information (Met, including P, PET, SD, and T) are all utilized to derive the release scheme. While other gain and loss terms in the reservoir water budget (e.g., water diversion, seepage, and evaporative loss) are unavailable for most reservoirs, the variables utilized in this chapter may contain information related to these factors. For example, reservoir evaporative loss is related to PET and water surface area, which in turn correlates with reservoir storage. Experiments 2 and 3 inputs exclude reservoir storage and inflow, respectively to evaluate the importance of reservoir

Table 2.1: Experiments using data-driven models with different time scale configurations and subsets of input variables, including inflow (*I*), storage (*S*), precipitation (*P*), potential evapotranspiration (*PET*), snow depth (*SD*), and air temperature (*T*).

Time Scale	Experiment	Training variables
Daily (D)	D-1	*I, S, Met (P, PET, SD, T)*
	D-2	*I, Met (P, PET, SD, T)*
	D-3	*S, Met (P, PET, SD, T)*
	D-4	*I, S*
	D-5	*I*
	D-6	*S*
Weekly-Daily (WD)	WD-1	*I, S, Met (P, PET, SD, T)*
	WD-2	*I, Met (P, PET, SD, T)*
	WD-3	*S, Met (P, PET, SD, T)*
	WD-4	*I, S*
	WD-5	*I*
	WD-6	*S*
Monthly-Daily (MD)	MD-1	*I, S, Met (P, PET, SD, T)*
	MD-2	*I, Met (P, PET, SD, T)*
	MD-3	*S, Met (P, PET, SD, T)*
	MD-4	*I, S*
	MD-5	*I*
	MD-6	*S*

information. Meteorological information is hidden in Experiment 4 to assess the impacts of meteorological forcing on reservoir release. Experiment 5 derives the release scheme only from the observed inflow records. Experiment 6 explores whether the actual storage alone is able to capture reservoir release decisions. It is noted that based on the specified subset of inputs, DDMs will further infer the importance of these variables on predicting reservoir releases via the training process. Results of these experiments will be used to guide further

sensitivity analysis based on models.

In all the experiments, this chapter uses the Long Short-Term Memory (LSTM, Hochreiter and Schmidhuber, 1997), as the DDM in each layer. As a powerful type of RNN, LSTM can learn temporal dependencies in both long and short terms and has a wide range of applications in hydrology and water resource management (Feng *et al.*, 2020; Kratzert *et al.*, 2018, 2019; Shen, 2018; Sit *et al.*, 2020; Xu and Liang, 2021; Yang *et al.*, 2021; Zhang *et al.*, 2018). The internal calculations of the LSTM cell are expressed as

$$\begin{aligned} i_t &= \sigma\left(W_{xi} \cdot x_t + W_{hi} \cdot h_{t-1} + b_i\right) \\ f_t &= \sigma\left(W_{xf} \cdot x_t + W_{hf} \cdot h_{t-1} + b_f\right) \\ g_t &= \tanh\left(W_{xg} \cdot x_t + W_{hg} \cdot h_{t-1} + b_g\right) \\ o_t &= \sigma\left(W_{xo} \cdot x_t + W_{ho} \cdot h_{t-1} + b_o\right) \\ c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \tag{2.2}$$

where W_{xi}, W_{xf}, W_{xg}, and W_{xo} are learnable weights of inputs x_t, W_{hi}, W_{hf}, W_{hg}, and W_{ho} are learnable weights of the previous hidden states , and b_i, b_f, b_g, and b_o are biases of the four gates, respectively. σ means sigmoid function, tanh is hyperbolic tangent function, and $\odot$ represents element-wise multiplication.

For the single-layer models (D1, $\cdots$, D6), the LSTM model is trained by minimizing the mean square error of daily release. For hierarchical time scale models (WD, MD), this chapter utilizes the iterative training strategy as mentioned in Section 2.2.1 to gain the optimal weights and bias. The two LSTMs are trained together by minimizing the mean square errors of reservoir release at both time scales, then the optimal parameters can be obtained by

$$\min_{\theta} \quad \frac{1}{T}\sum_{t}^{T}\left(y_t^{(1)} - \hat{y}_t^{(1)}\right)^2 + \frac{1}{T}\sum_{t}^{T}\left(y_t^{(2)} - \hat{y}_t^{(2)}\right)^2 + \frac{1}{T}\sum_{t}^{T}\left(y_t - \hat{y}_t\right)^2 \tag{2.3}$$

where $y_t^{(1)}$ and $\hat{y}_t{}^{(1)}$ are the observed and simulated release at the monthly/weekly scales, $y_t^{(2)}$ and $\hat{y}_t{}^{(2)}$ are the observed and simulated release deviations at the daily scale, y_t and $\hat{y}_t$ are the observed and simulated release at the daily scale, θ represents the neural network parameters. The data at the coarse scale is remapped to the daily scale by resampling to ensure consistent lengths of data at both coarse and daily scales. Sixty percent of time series data are used during the training process, 10% of them for validation, and the rest for testing. The Adam optimizer (Kingma and Ba, 2014) is applied for primary training and Stochastic Gradient Descent (SGD, Robbins and Monro, 1951) for finetuning. The number of training epochs and number of hidden units are found through trial-and-error. The learning rate during the pretraining process is 10^{-4} to 10^{-5} and the number of training epochs does not exceed 100, while the learning rate schedule is more complex during the finetuning process. Early stopping is implemented to decrease the probability of overfitting. To ensure the fairness of subsequent comparisons, the total number of parameters for both single-layer (D) and hierarchical time scale models (WD, MD) is constrained to be identical. Specifically, the hidden size in the single-layer model is almost equivalent to the sum of hidden size in all DDMs in the two-layer model. Concretely, this chapter sets the hidden size of daily single models for all reservoirs as 10, 15 or 20 to avoid excessively complex DDM models, ensuring that the maximum total number of parameters in single and hierarchical models does not exceed 2,000. The hidden size in the first layer of the hierarchical models is 5, 10 or 15, and that in the second layer is correspondingly adjusted. The Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) of daily reservoir release is used for assessing model performance in all experiments. To mitigate random effects arising during training, the author initializes and train the models with different random seeds, calculating average performance metrics across the five trials. All the performances mentioned in the following sections are NSEs evaluated on the test sets. It is noted that the multi-layer configuration is flexible to use other data-driven algorithms.

2.3 Results

2.3.1 Performance of DDMs with Various Time Scale Configurations and Input Variable Combinations

Results from the three groups of experiments revealed noticeable differences in reservoir release simulation accuracy when the models use various time scale configurations and combinations of input variables (Figure 2.2). In experiments employing the same training variables, DDMs at the daily scale are capable of simulating the dynamics of reservoir release, and the two-layer hierarchical model (WD) exhibits consistent superiority over the daily model (D) in terms of accuracy, as evidenced by the probability of NSE exceedance across all reservoirs (Figure 2.2). MD configuration proves capable of outperforming the daily single scale model in select cases, notably for the majority of reservoirs examined in Experiments 2, 3, 5, and 6. In Experiment 1 with the most comprehensive input data set, the median NSE for all reservoirs is 0.900, 0.831, and 0.872 for WD, MD, and daily configuration, respectively. The WD configuration achieves NSE higher than 0.8 in more than 88% of reservoirs, compared to 61% and 77% for the MD and D configurations, respectively. The WD configuration generally outperformed the MD configuration in most experiments. This may be attributed to the fact that weekly scale data provides four times more information than monthly scale data, thereby enabling the DDMs to be trained on more samples, even though both are resampled to the daily scale. Additionally, the finer resolution of the weekly scale may more accurately capture the variability of release decisions compared to the coarser monthly scale.

For all time scale configurations, reservoir inflow and storage are two key explanatory variables for modeling release behavior in most reservoirs, as indicated by the marginal performance gap between Experiments 1 and 4. With only reservoir inflow as model input in Experiment 5 (Figure 2.2e), the median NSE reaches 0.655, 0.826, and 0.762 for daily,

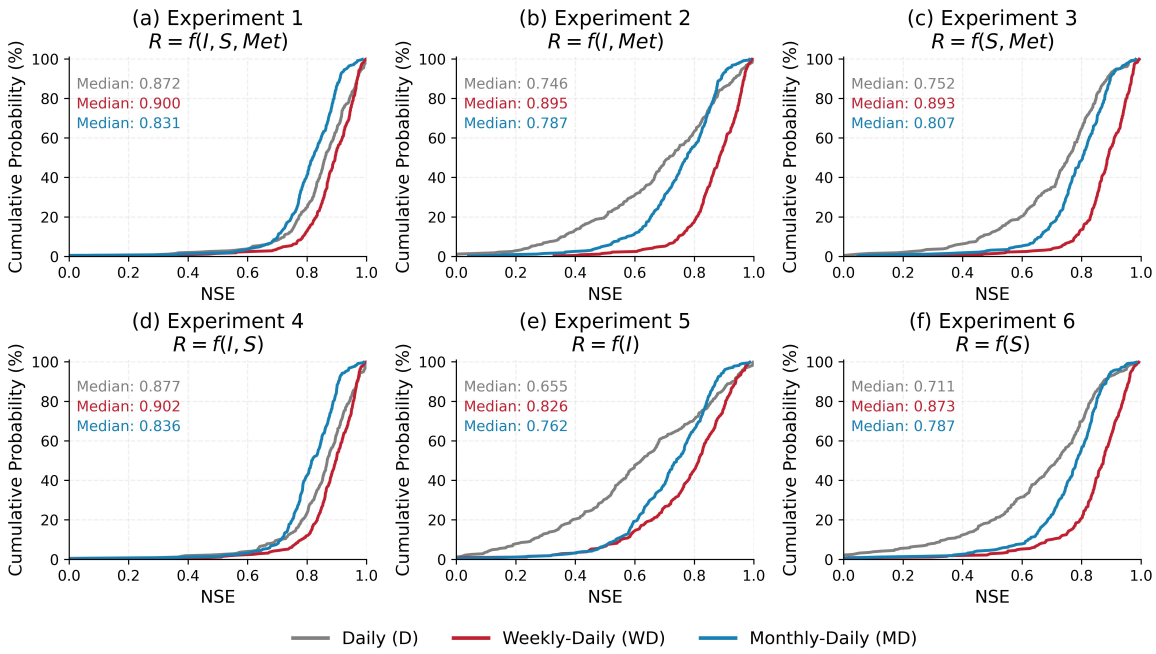

Figure 2.2: Probability of exceedance of Nash-Sutcliffe Efficiency (NSE) for all reservoirs resulting from single and hierarchical time scale models with different decision variables (Table 2.1).

WD, and MD temporal configuration, respectively. The inflow provides the most predictive power in reservoirs with relatively small storage and/or navigation purpose, particularly for run-of-river reservoirs located along the Columbia River or the Arkansas River, where there is a strong linear relationship between inflow and release at daily scale and the impact of storage can be negligible. Although the inflow-only models in Experiment 5 do not explicitly consider reservoir states, the LSTM architecture is expected to use the "hidden state" and "cell memory" to store accumulated inflow as a proxy for reservoir storage trend and use this information to simulate reservoir releases. However, due to the lack of other reservoir water budget terms such as water diversion, seepage and evaporative loss, the accumulated inflow cannot fully replace reservoir storage. Therefore, it is not ideal for a single time scale DDM to simulate the state of a reservoir system without storage as an important constraint, especially for reservoirs in the west mountainous regions usually

designed for water supply and hydropower generation. Because reservoir storage is closely related to operational purposes, and its seasonal variations typically reflect the consequences of human interventions on the natural system, storage volume (or water level) is strongly recommended as an independent variable input into the reservoir operation model.

The DDMs with storage alone as input in Experiment 6 have slightly higher predictive power compared to inflow-only models in Experiment 5 (Figure 2.2f) and produce median NSE of 0.711, 0.873, and 0.787 for Daily, WD, and MD configuration, respectively. Using storage as the model input captures operations of reservoirs with relatively large storage capacity and/or reservoirs with water supply purpose where the release largely depends on the reservoir water level. In addition, reservoir storage serves as a proxy for reservoir water level and water surface area (both can be retrieved from the reservoir characteristic curve). The reservoir storage together with PET may implicitly contain information regarding reservoir evaporative loss, which is important in arid and semi-arid regions. Storage-release rule curves are commonly used by reservoir operators (Yang *et al.*, 2016), which cover the seasonal patterns of reservoir operation but the interannual variability of inflow is likely missing in such curves. At a monthly or seasonal scale, water control plans designed for specific purposes or hydroclimatic conditions that influence the upstream flow rate may exhibit low year to year variation within decades. At daily or sub-daily scale, however, reservoir inflow can vary a lot due to emergency events or weather fluctuations, especially for those reservoirs with complicated operational conflicts between multiple objectives or climate-sensitive reservoirs (such as reservoirs in the New England regions faced with potentially increasing flooding risks under the context of global warming). Although actual rule curves implemented by reservoir operators could provide substantial information to understand the decision-making process of water resource management, it does not adequately represent the operation tradeoffs under various inflow conditions. Reservoir inflow should be considered as a paramount input while building data-driven operation

models. Combining the inflow and storage in Experiment 4, the median NSE improves to 0.877, 0.902, and 0.836 for daily, WD, and MD temporal configuration, respectively.

The performance improvement from including hydroclimatic variables (e.g., P, PET, SD, and T) is investigated by comparing accuracies of DDMs in Experiment 1 versus 4, Experiment 2 versus 5, and Experiment 3 versus 6. When both inflow and storage are used (Experiment 1 vs. 4), the improvement from additional hydroclimatic forcing is negligible (mean NSEs increase no more than 0.05). For DDMs with only inflow (Experiment 2 vs. 5) or storage (Experiment 3 vs. 6), adding hydroclimatic information slightly enhances the model performance, which is not unexpected as DDMs typically benefit from more input information. Nevertheless, it may also underscore the potential of incorporating hydroclimatic conditions in reservoir release modeling (Denaro *et al.*, 2017), particularly in regions where reservoir operation records are scarce.

2.3.2 Effect of DDMs Hierarchical Temporal Configuration on Capturing Reservoir Operation Behavior

Figure 2.3 further illustrates the improvement of the hierarchical framework for reservoir operation modeling and the nuances of such improvement with/without hydroclimatic information at different time scales. Hierarchical temporal scale models work for some cases, although they do not always perform better than the models constructed on the single time scale under the same experiment settings. When one of the dominant explanatory variables (e.g., inflow or storage) is missing, a better organization (i.e., hierarchical temporal configuration) of the explanatory variables further enhances the performance. For example, in Experiments 2, 3, 5, and 6, more than 60% of reservoirs benefit from re-arranging the training data in hierarchical configuration (WD and MD) compared to the single daily scale configuration, although the DDMs in this experiment contain the same amount of information. This highlights the benefits of incorporating the multi-temporal scale of

reservoir behaviors into the configuration of DDM to capture the reservoir operation under various targets, in particular when hydrometeorological information or reservoir operational records are limited.

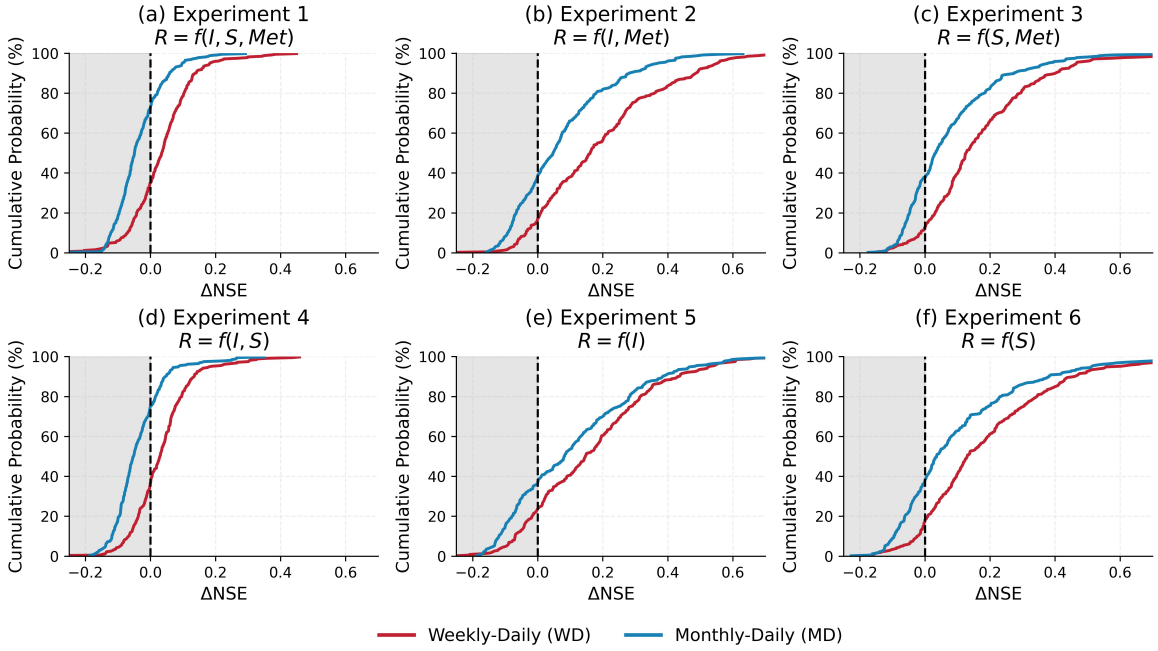

Figure 2.3: Improvement of Nash-Sutcliffe Efficiency (NSE) by hierarchical time scale framework ($\Delta\text{NSE} = \text{NSE}_{\text{hierarchical}} - \text{NSE}_{\text{single}}$) in (a) Experiment 1, (b) Experiment 2, (c) Experiment 3, (d) Experiment 4, (e) Experiment 5, and (f) Experiment 6. $\text{NSE}_{\text{hierarchical}}$ represents the performances of hierarchical time scale models (WD and MD), while the $\text{NSE}_{\text{single}}$ is the performance of a single time scale model (D).

Regardless of the experimental settings, WD consistently outperforms another two-layer hierarchical model MD in simulating reservoir release decisions. Specifically, in Experiment 6 (Figure 2.3f) with the reservoir storage only as model inputs, performances of about 80% of reservoirs have been improved by the hierarchical framework (WD), and it is more prominent than the MD where the first layer simulates the reservoir release on the monthly scale. It probably indicates that sub-monthly operational information and hydroclimatic forcing, which shows significant short-term variability, may provide a substantial portion of the information needed for accurate reservoir operation modeling. By incorporating

information on moderate and fine time scales, WD DDM can well capture the complex dynamics of reservoir operations and yield highly accurate predictions, which may help inform the development of more effective and efficient reservoir management strategies in the face of increasing hydroclimatic variability.

2.3.3 Spatial Pattern of DDM Reservoir Operation under Various Temporal Configurations

Figure 2.4 shows the spatial distribution of average NSE improvement by WD and MD from Daily configuration for all six experiments, respectively. When the dominant explanatory variables (i.e., inflow and storage) are fed as model inputs (Experiments 1 and 4), most reservoirs across the CONUS do not benefit significantly from the hierarchical temporal scale framework (Figures 2.4a and 2.4d). This can be attributed to the fact that the daily single model performs well for most reservoirs with a median NSE higher than 0.85 (Figures 2.2a and 2.2d), which demonstrates the efficacy of DDMs for reservoir release simulations. When the most relevant variables are sufficiently represented in the data, additional methods for regulated flow simulation refinement may not be necessary. Hierarchical models face challenges in improving the accuracy of models for reservoirs that primarily serve a single purpose or are predominantly operated at a single time scale. For instance, the hierarchical time scale model does not improve and even degrades the release modeling of run-of-river reservoirs. In the New England district, where many reservoirs have limited storage capacity and are primarily used for flood control during flood seasons and recreation during non-flood seasons, hierarchical models are less effective across all experiments (Figure 2.4). This highlights the importance of identifying the appropriate modeling resolution to match the time scale at which reservoir release decisions are made.

Hierarchical models send positive signals for reservoirs in the Midwest. The hierarchical DDM improves NSE over Daily scale in many reservoirs in the western United States as

Figure 2.4: Spatial distribution of average Nash-Sutcliffe Efficiency (NSE) improvement ΔNSE from Daily scale to hierarchical time scale configuration of Data-driven models in (a) Experiment 1, (b) Experiment 2, (c) Experiment 3, (d) Experiment 4, (e) Experiment 5, and (f) Experiment 6. The circles with black solid edges represent reservoirs labeled as "lock and dam."

shown in Figure 2.4, and the magnitude of improvement in model performance varies across different experiment setups. For Experiment 1 and 4 that includes both inflow and storage as model inputs, the average NSE improvement ΔNSE is subtle (about 0.1 0.25) for some reservoirs in Montana, Utah, New Mexico and Texas (Figures 2.4a and 2.4d). It implies that the hierarchical model is effective in capturing reservoir release behavior in western regions, at least to a comparable degree as the daily single model. For Experiment 2 and 5 that do not contain storage fed into models, release simulations are boosted by hierarchical temporal scale framework for reservoirs on the High Plains (e.g., Texas, Oklahoma, Kansas), highlighting the signature of seasonal cycle of water supply operation in these reservoirs. Reservoirs that provide water for agricultural irrigation or municipal/industrial use often base release decisions on the water level or storage status. In situations where operational records of storage are unavailable, comprehensive utilization of inflow data across various temporal scales may serve as a compensatory mechanism. Many water-supply reservoirs maintain nearly constant storage volume at the start and end of an operational year, resulting in a nearly balanced inflow and release volume at a certain temporal scale (monthly, seasonally, or annually). Thus, it becomes feasible to detect reservoir behavior when changes in inflow over the preceding months or weeks are known. It would facilitate accurate estimates of regulated flow regimes in the absence of readily available data sets on water level or storage under future scenarios. In the case of reservoirs located in the Rocky Mountains and the Colorado River basin, the hierarchical model consistently enhances the accuracy of release modeling, regardless of whether inflow or storage is excluded as explanatory variables (Experiments 2 and 5; Experiments 3 and 6). As stated in Section 2.3.1, inflow generally reflects short-term variability or the effects of fine-scale weather fluctuations, while storage represents the cumulative hydrologic response during past periods. The absence of either of these dominant factors results in a loss of vital information for accurate release modeling. Hence, the behavior of reservoirs in the west cannot be fully captured by DDMs

at a single temporal scale. Among the observational records analyzed in this chapter, 216 reservoirs serve at least three purposes and 79 out of 327 serve at least five purposes. In spite of accounting for multiple time scales may not be imperative for simpler reservoirs that serve fewer purposes or operate under less complex conditions, it is crucial for effectively modeling multi-purpose reservoirs and multi-reservoir systems.

In summary, the analysis indicates that reservoir release modeling can be enhanced by leveraging the availability of adequate information, with particular emphasis on key explanatory variables. The inclusion of meteorological forcing data may also be beneficial for accurate simulation. In situations where the records of reservoir inflow or storage are inaccessible, the comprehensive utilization of multiple temporal scales can lead to improved modeling outcomes.

2.3.4 Dominant Variables of Reservoir Release across Time Scales

Although DDMs frequently achieve remarkable results in model performance, further sensitivity analysis would help to diagnose and interpret the empirical relations captured by the "black-box" DDMs. Different DDMs have individual strengths and weaknesses in simulating the reservoir release, and few single models could consistently outperform others (Yang *et al.*, 2021). Performances of different DDMs can vary widely by the modeling schemes, by the ways of training data structure, as well as by the statistical measurement used. Model interpretability benefits further improvement in performance and providing insights on anthropogenic behaviors under hydroclimatic variabilities. The hierarchical configurations of DDMs allow us to explore whether reservoir operation depends on different variables and how the dominant variables change across time scales, thus providing an interpretable avenue to enhance the understanding of reservoir behavior.

A prevalent method for enhancing interpretability is to analyze variable importance. Many approaches can be taken to assess feature importance of machine learning models.

Wei *et al.* (2015) conducted a comprehensive review of various techniques for variable importance analysis in different disciplines and analyze their relative merits. Recently, Quinn *et al.* (2019) used time-varying sensitivity analysis to open the black box of multi-reservoir operation models. Additionally, Shapley Additive Explanations (Lundberg and Lee, 2017) and permutation feature importance (Breiman, 2001; Fisher *et al.*, 2019) have gained popularity in recent years. This chapter uses well-trained DDMs to conduct a variable importance analysis that explores the impact of decision variables on reservoir release schemes across different time scales. The author employes the permutation feature importance method to measure variable importance, which involves randomly permuting feature values in the input data and examining the effect on model performance, as measured by a specific metric (such as NSE in this study). The extent of the decrease in performance reflects the relative importance of the feature, with a greater decline in performance indicating a more influential feature in the model. Then the variable that leads to the largest change is referred to as the most important variable, or dominant factor.

Figure 2.5 displays the most important variable for each reservoir across CONUS on the different time scales (daily, weekly, and monthly) of Daily, WD, and MD configurations in Experiment 1 that contains all the variables (inflow, storage, precipitation, *PET*, *SD*, and air temperature) as model inputs. For half of reservoirs (163 out of 327), the same variable has critical influences on the release on all time scales (daily, WD-weekly, MD-monthly, WD-daily, MD-daily), likely implying the consistency of their operating strategies and trade-offs on various time scales, and there may be a primary purpose that dominates the operation process throughout the year. For 120 of these reservoirs, inflow plays a decisive role in reservoir release at all time scales, while storage volume is the most instructive variable for 42 of these. Daily models with good performance (e.g., reservoirs labeled as "lock and dam" along the Arkansas River and Columbia River) generally identify inflow as the primary variable, as inflow exhibits high short-term variability and can effectively

inform the daily release decision. Reservoirs located on the High Plains, where water level is a crucial factor in release operations, consistently show storage as the dominant factor influencing release decisions. The findings of variable importance in California and the High Plains differ slightly from those reported by Hejazi *et al.* (2008) (e.g., many reservoirs in these two regions reported by Hejazi *et al.*, 2008 do not consider storage as the dominant factor), who investigated the dependency of operators' release decisions using the method of information theory based on weekly/monthly operational records. It should be noted that Hejazi *et al.* (2008) included past release as a decision variable, while this chapter did not consider it as a model input. Furthermore, the operational data set utilized in this chapter is updated to 2016, which may account for this discrepancy. Martis Creek Lake, located in the Sierra Nevada Mountains outside the town of Truckee, serves the dual purpose of flood control and recreation, with precipitation (P) being the most predictive variable for reservoir inflow at all timescales. The lake is situated in a headwater watershed with a small contributing area, which further supports the use of P as a reliable proxy for inflow prediction. It is worth mentioning that for two reservoirs, the Elephant Butte Reservoir in New Mexico and the Moon Lake in Utah, PET has a major effect on reservoir release at the daily, WD-weekly, and MD-monthly scales (maps along the diagonal shown in Figure 2.5), which could involve considerable reservoir evaporation and water use for agricultural irrigation in the arid, semi-arid western mountains. These results of model-based sensitivity analysis further validate the findings given by the comparison of Experiments 1 and 4. That is, reservoir inflow or storage volume has a paramount influence on the release decision rather than hydroclimatic forcing. Only for very few reservoirs, hydroclimatic forcing directly dominates the reservoir release.

The most important feature on the release decisions can extend to aligning with their functional roles and have operational implications. For instance, at the daily scale, the inflow-dependent reservoir (along the Arkansas River, Columbia River, and coastal areas,

see the first panel in Figure 2.5), typically serve the purposes of flood control, run-of-the-river hydropower generation, etc. For these reservoirs, the accuracy of inflow forecasts could substantially influence release decisions. The storage-dependent reservoirs (mainly in the high plains and Midwest), which are used for water supply, irrigation, impoundment hydropower, etc., suggest the storage capacity could greatly influence the downstream flow conditions simulation.

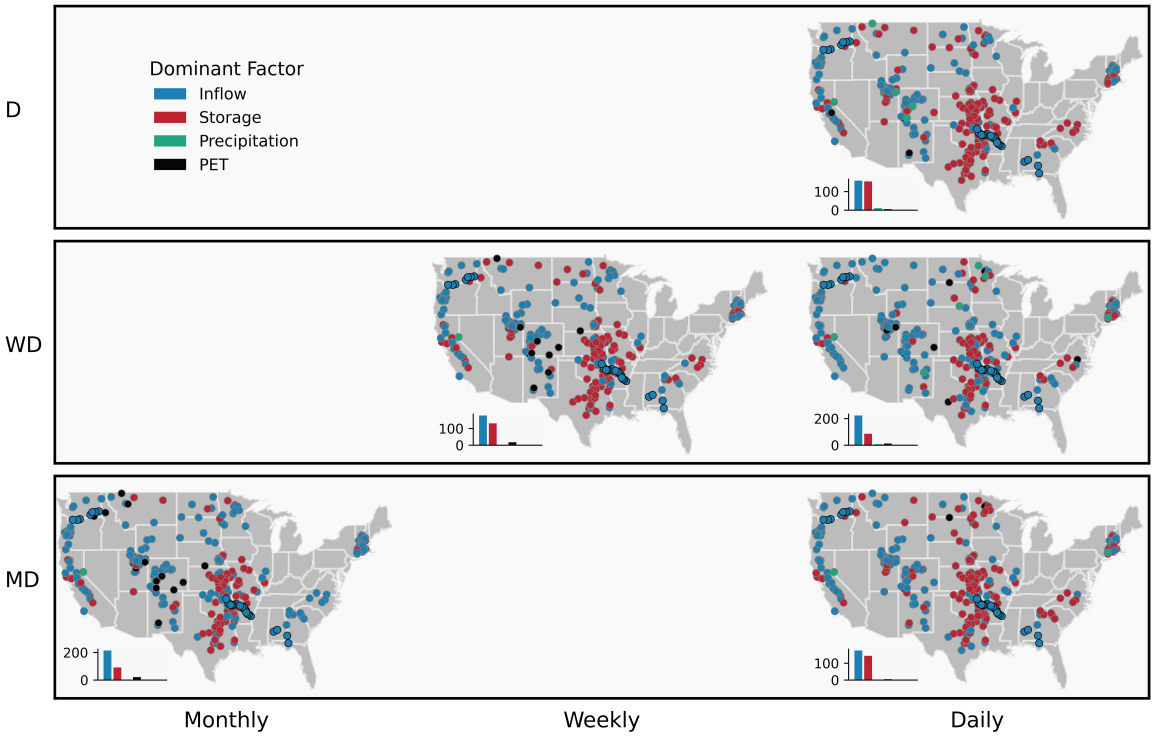

Figure 2.5: Spatial distribution of dominant factors across daily, weekly, and monthly scales. The circles with black solid edges represent reservoirs labeled as "lock and dam." The inset at the bottom left corner depicts the number of reservoirs in which a certain variable (inflow, storage, precipitation or potential evapotranspiration) is identified as the dominant factor influencing release decisions.

It is interesting to notice that more than one third of (117 out of 327 for WD; 108 out of 327 for MD configuration) reservoirs vary in their dependency on decision variables at different time scales (shifted from weekly to daily in the WD; from monthly to daily in the MD configurations), suggesting that reservoir operators consider different information at

different time scales to fulfill multiple designed purposes. For MD shown in Figure 2.5, at the monthly scale, operations of 214 reservoirs primarily depend on the reservoir inflow, and 91 reservoirs rely more on storage volume. At the daily scale, the number of reservoirs with major dependency on inflow decreases to 175 and that of reservoirs relying more on storage volume increases to 174. From the coarse scale to the fine scale, nearly 20% of reservoirs (64 out of 327) shift their primary dependence from inflow to storage volume. As mentioned in Section 2.3.3, many reservoirs tend to maintain nearly constant water level or storage at the beginning and end of an operational year, which can result in a balance between the total volume of inflow and release at certain time scales (e.g., annually, seasonally). Consequently, it is not surprising that for almost two-thirds of the reservoirs studied, inflow exhibits the strongest relation with release at the monthly scale. At the daily scale, operators tend to give greater weight to current or recent storage status (or water levels) when making release decisions, since reservoir storage is a crucial factor in determining the availability of water for downstream users or for maintaining water levels within acceptable limits. Although neither *SD* nor temperature is detected as the dominant factor at any of the reservoirs, it would be premature to dismiss these two factors as unimportant. This is probably on the ground that the variable importance analysis used in this chapter is model-based rather than based on observational data, which sometimes might produce misleading results due to inadequate feature selection or inappropriate model configuration, particularly when *SD* or air temperature is tightly linked to other explanatory variables such as *P* or *PET*. It is important to exercise caution in interpreting the results. Additionally, this chapter merely focuses on the most important variables in this chapter, but *SD* and air temperature are likely to play a substantial role in snow-dominated, high-altitude mountain reservoirs.

2.3.5 Reservoir Release Behaviors across Time Scales

Compared to attempts to capture reservoir operation at a fixed time scale, the hierarchical temporal configuration in this chapter demonstrates improved model performance while utilizing the same input information, particularly when essential decision variables such as inflow or storage are inaccessible. In addition, the sensitivity analysis suggests that operation in many reservoirs depends on different information at different time scales. In the following paragraphs, the author picks the multi-purpose Belton Lake reservoir to elaborate how various operation targets manifest their signatures at different time scales, thus requiring hierarchical temporal configuration to fully capture the tradeoffs among multiple operation targets.

The Belton Lake (TX00002) is located on Leon River in Texas with 536.8 million cubic meters (or 435,500 acer-feet) conservation capacity (Texas Water Development Board, 2015) and the maximum storage volume of around 1,440 million cubic meters. The 192-feet high dam maintains the water level at elevation between the conservation pool elevation of 594 feet and the crest elevation of 631 feet, with flood control, water supply and irrigation as listed operation targets under the management of the U.S. Army Corps of Engineers. The annual mean inflow volume is 641.5 million cubic meters. Belton Lake provides an example with a large storage capacity in a humid subtropical climate. The DDM in Experiment 5 (with inflow only) has NSE of 0.848, 0.969, 0.920 for Daily, WD and MD configuration, respectively. The DDM identifies reservoir storage as the dominant variable on release at Daily, WD, and MD scales, respectively.

Figure 2.6 shows the scatter plots of release versus inflow and storage versus inflow at various time scales. At the annual time scale (Figure 2.6d), the outflow is highly correlated with inflow, suggesting the reservoir has seasonal flow regulating capacity. The slightly lower annual release than the inflow (Figure 2.6d) indicates water balance is roughly held

on annual time average. Water supply withdrawals made through pumping or diversion have a limited impact on the mass balance. The randomness between monthly inflow and release (Figure 2.6c) shows a wide range during different seasons indicating the seasonal buffering capacity of the reservoir storage. The storage versus release scatter plot shows reconcilable patterns starting from the monthly scale.

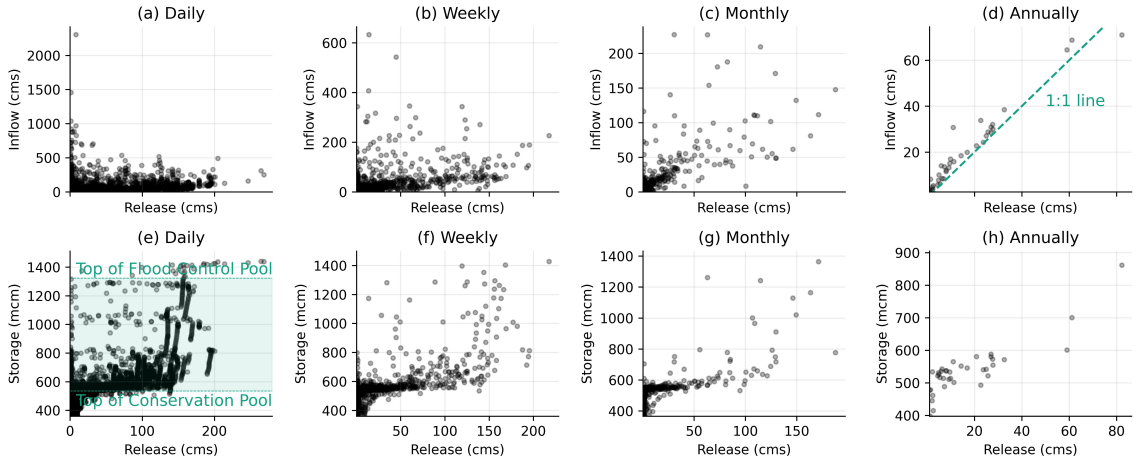

Figure 2.6: Relationship between inflow and release at (a) daily, (b) weekly, (c) monthly, and (d) annual scale; Relationship between reservoir storage and release at (e) daily, (f) weekly, (g) monthly, and (h) annual scale of Belton Lake (TX00002).

Figure 2.7a shows the flow duration curves of Belton Lake inflow and releases simulated by different DDM configurations. The Daily, WD, and MD achieve similar predictability to capture the regulation during medium to high flow conditions (i.e., flow larger than 20% exceedance probability). The Daily scale DDM overestimates the low to medium flow range (i.e., flow less than 40% exceedance probability) with given inflow only, and the MD scale DDM slightly overestimates the medium flow (i.e., flow between 25% and 45% exceedance) and underestimates the low flow range (i.e., flow less than 60% exceedance probability). The WD scale DDM reproduces the flow duration curve (FDC) for almost all flow conditions although not perfectly.

The hydrograph of Year 2002 in Figure 2.7b shows the seasonal pattern and short-term

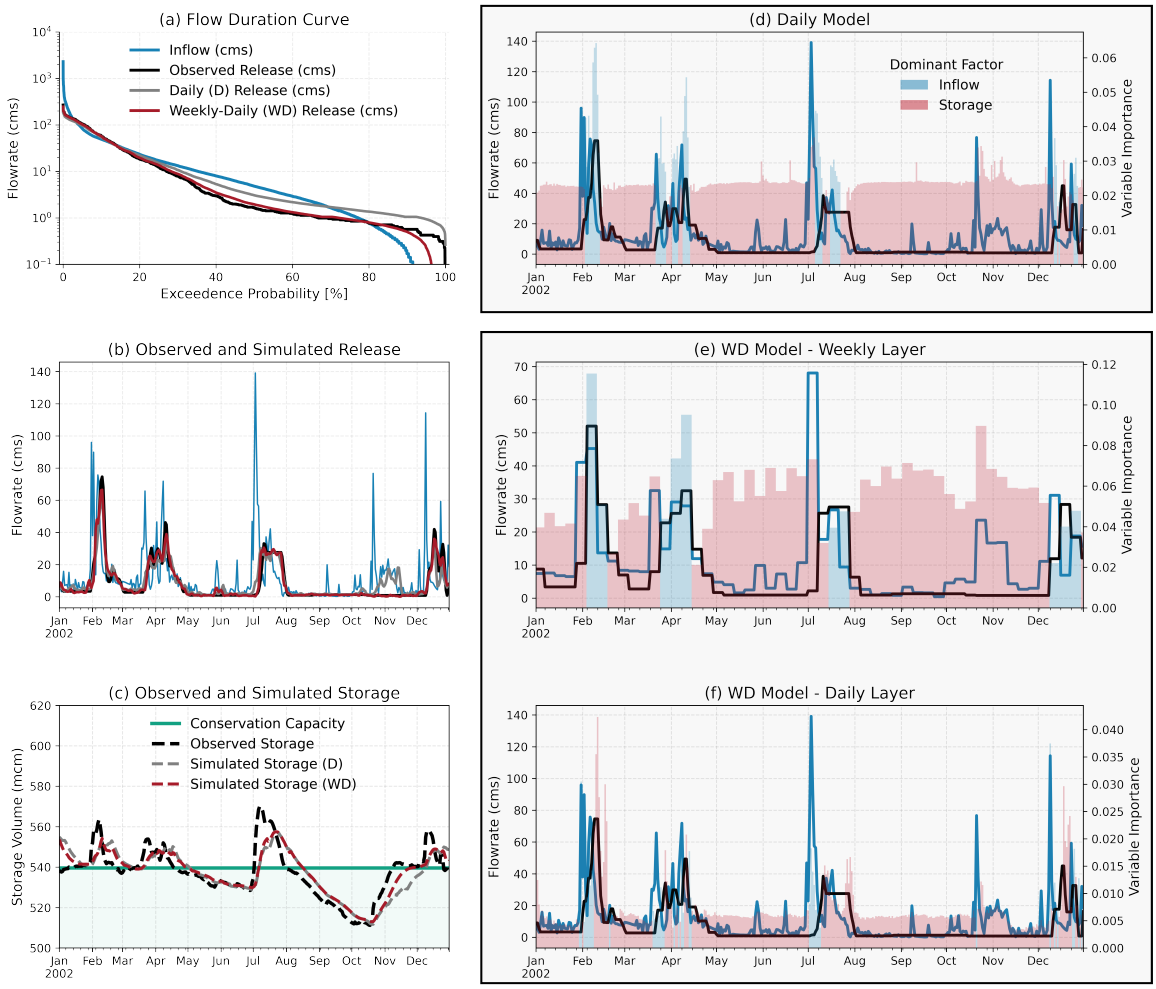

Figure 2.7: Experiments exploring the dominant factors and simulated outputs of Belton Lake (TX00002) in Texas. Comparisons of observed and simulated release in Experiment 5 (only inflow as inputs) shown in panel (a) Flow Duration Curve and (b) hydrograph during the calendar year 2002 (test period); (c) Conservation capacity, observed storage, and simulated storage during the calendar year 2002, where the simulated storage is derived from the water balance equation, inflow, and simulated release; Time-varying dominant factors from Experiment 4 (both inflow and storage as model inputs) shown in (d) daily model, (e) weekly layer of WD model, and (f) daily layer of WD model during the calendar year 2002. The sub-axis presents the computed absolute variable importance values obtained through the application of Shapley Additive Explanations method (Lundberg and Lee, 2017) The dominant factors, characterized by the highest absolute variable importance values, are denoted by different colors.

variation produced by different DDM configurations. The Daily Scale DDM tends to exhibit a faster decay in release following flood events, since the daily scale model is sensitive to the daily input and lacks long-term information. The WD scale configuration demonstrates superior performance in capturing both seasonal water supply and flood control release at Belton Lake. As an illustration, in November 2002, when the daily model produces a false release response while the hierarchical models do not. It exposes the shortcomings of a single daily scale model for multi-purpose reservoirs that consider reservoir storage as an influential factor. Many large reservoirs in Texas adhere to a general strategy based on minimizing the risk and consequences of releases contributing to downstream flooding in the flood seasons, while ensuring the maximum design water surface is never exceeded (as shown in Figure 2.6e). Release decisions are contingent upon the flood control pool storage capacity. In the non-flood seasons, these reservoirs strive to maximize water levels within the conservation pool, without surpassing its upper limit (i.e., the top of conservation pool). When storage information is not explicitly provided as an input, it can be challenging for a daily single model to consistently and accurately respond to inflow information. Although DDMs are expected to derive storage information from the physical constraints (e.g., water balance equation) and the accurately simulated release time series (Figure 2.7c), challenges remain due to the inaccurate simulation in release, error accumulations, missing water budget terms, etc. If only reservoir inflow is given, which typically represents short-term hydrologic variability, the long-term target may be overlooked by a daily single time scale model due to the absence of long-term hydrologic indicators. An example from Figure 2.7c illustrates that from September to November 2002, the water level (storage status) fell below the target conservation capacity, resulting in the reservoir not releasing water in response to inflow events during this period. Storage is recognized as the dominant factor that determines the reservoir release decision at both daily and weekly scales during this period (Figures 7d and 7e). In the absence of storage that can reflect long-term hydrologic variability, the Daily

model fails to capture the implicit long-term patterns inherently embedded in the absent key variable. This "short sight" explains the erroneous response observed in Figure 2.7b (gray line), further emphasizing the importance of fully utilizing multiple temporal scales of information. In contrast, the hierarchical model with multiple time scales can better incorporate the complex dynamics of the reservoir system, which can lead to more reliable and robust simulation results.

These observations highlight the importance of appropriately organizing training data at various time scales in order to enable machine learning techniques to capture the underlying relationships inherent at each time scale. This chapter also uses other machine learning techniques (e.g., Random Forest) to configure the hierarchical DDM and achieved satisfactory results, suggesting the predictability is not limited by the choice of specific machine learning models. From the perspective of effectively training the machine learning models, hierarchical temporal configuration not only yields better predictability, but also provides more meaningful interpretations of the DDM.

2.4 Discussion

2.4.1 Strategies and Limitations of Data-Driven Reservoir Operation Modeling

This chapter employs LSTM networks to simulate reservoir release decisions, primarily due to their similarities to traditional hydrological models to some extent—for example, current hydrological fluxes are determined by current inputs and past states. The strength of LSTM networks lies in their ability to learn nonlinear patterns and long-term dependencies, making them ideal for simulating reservoirs where the hydrological behavior may change over time. LSTMs are expected to be suitable for modeling when the decision variables (or model inputs) exhibit temporal dependence. While LSTM networks have become widely used in the hydrology community, barriers may exist due to the requirement of a large

amount of training data and careful finetuning processes to achieve accurate results. In addition, the measurement of feature importance in neural networks is not so straightforward and makes it lack interpretability. It is essential to acknowledge that LSTM networks may not be the optimal choice for simulating reservoir operations all the time, especially in cases where actual operation rules are explicit. For instance, in some highly engineered watersheds in the western United States, which are equipped with a large number of dams, the reservoir release patterns can deviate considerably from the natural flow characteristics of the system. These deviations are a result of the complex interactions between the reservoir operations and the hydrological processes, which can be influenced by a range of factors such as climate change, water demand, and land use change. In these cases, other white-box models such as Classification and Regression-Tree or RF, which are more intuitive for decision-makers and excel in capturing patterns from various features, may be more appropriate (e.g., Yang *et al.*, 2016). Moreover, a notable drawback of LSTM and other RNN-based models typically pertains to their dependence on data continuity, particularly when the lookback or look forward window is extensive. For instance, in the context of rainfall-runoff or models involving surface-groundwater interaction, such a window may span as much as 180, 270, or 365 days (e.g., Kratzert *et al.*, 2019). While preprocessing techniques can handle missing data to create a continuous time series as inputs, the usefulness of models needing continuous data might be limited in situations where reservoir operation records are scarce.

Unlike many well-established DDMs for reservoir operations, such as those developed by Turner *et al.* (2021); Chen *et al.* (2022a,b); Dong *et al.* (2023); Brunner and Naveau (2023), this chapter omits reservoir storage simulation, which is a frequently pursued research objective in the development of reservoir operation models. It is because this chapter aims at investigating the significance of multi-timescale information in data-driven reservoir operation modeling. Specifically, this chapter seeks to examine the impact of incorporating multiple temporal scales of decision variables in the construction of models

for reservoir release and aspires to contribute to the ongoing effort to enhance the performance and robustness of data-driven regulated flow simulations. It is noteworthy that the interdependency between reservoir inflow, storage, and release across various time scales (as pointed out in Section 2.3.3 and the example shown in Figure 2.6) can be leveraged to extract informative features for input into white-box models (i.e., feature engineering considering multiple scale temporal information), with the potential to enhance the balance between model performance and interpretability. By exploiting the rich temporal dynamics of reservoir operations data, it can facilitate a more comprehensive and interpretable representation of the underlying processes.

The feasibility of data-driven reservoir simulations can be further boosted through the use of hybrid strategies that combine rule-based or conceptual operation schemes with machine learning techniques (Dong *et al.*, 2023; Gangrade *et al.*, 2022). By leveraging expert knowledge in the form of appropriate feature engineering (Yang *et al.*, 2016, 2017), and by incorporating reservoir storage dynamics derived from a range of advanced sensing techniques (Chen *et al.*, 2022a,b; Eilander *et al.*, 2014; Sorkhabi *et al.*, 2022; Van Den Hoek *et al.*, 2019), it is possible to use DDMs to better reconstruct downstream flow in data-sparse regions, using meteorological forcing and inflow generated by hydrological models.

2.4.2 Hierarchical Nature of Anthropogenic Decisions

DDMs are generally not constrained by the complexity of the training data set and can achieve better prediction with more training variables. However, the results illustrated in Figure 2.4 suggest that in an identical experimental setup, employing congruent variables and model architecture while maintaining consistent model complexity (as indicated by an equivalent total parameter count), the hierarchical timescale model—which encompasses both coarse and fine scales and is anticipated to acquire an augmented amount of temporal data—does not invariably surpass the performance of a single-timescale model. It indicates

that reservoir operation decisions under different operation targets are associated with different time scales and require different information. Therefore, simply including more variables into the training data sets or increasing the hierarchical layers does not guarantee better predictability. This observation highlights the importance of providing appropriate information that matches the temporal resolution to capture reservoir release behavior under various targets.

Although the scaling issue in hydrologic processes has been well recognized by the hydrologic community, there are few studies to investigate the scaling of decision making in water resources management. In representing anthropogenic components (by either simulation or optimization approach) in hydrologic models, the decision makings are generally based on one single time scale. For example, farmers' irrigation decisions depend on soil moisture conditions. The reservoir operation policy is optimized to balance the tradeoff between water supply benefits and flood risk based on daily streamflow. The hierarchical temporal scale configuration of DDM in this chapter explicitly shows that the single temporal scale model cannot fully capture the reservoir release under various operation targets. Different operation targets are associated with different temporal scales and require corresponding hydroclimatic information. For example, the reservoirs in the Colorado River Basin use the seasonal snowpack condition to forecast the water supply (Bureau of Reclamation, 2022; Xiao *et al.*, 2018), while the hydroelectric generation is based on hourly demands from power grids.

Besides the dependence on cross-scale information, anthropogenic decisions also interact at different scales. Short-term decisions (e.g., operation of water resources infrastructure) are constrained by long-term decisions (e.g., planning of water resources infrastructure), and the objectives of decisions at different scales may require tradeoffs. For example, given the same amount of agricultural water supply, farmers can tradeoff between crop type and irrigated area (decisions made before growing season) and the actual irrigation intensity

(decisions made during growing season), which results in different water release amount and frequency. The hierarchical temporal configuration of DDM in this chapter recognizes the cross-scale interaction feature and handles this feature by simulating the daily release deviation from the weekly/monthly release. For traditional optimization formulation in water resources management, the author believes the hierarchical optimization (Karsanina and Gerke, 2018; Yeo *et al.*, 2007) would be a promising configuration to represent interactions of decisions made across scales.

As hydrologic models and observations continue to improve and provide better prediction, the ultimate question is how hydrologic prediction (and what types of prediction) can be effectively utilized to improve the operation of reservoirs. There are efforts to forecast informed reservoir operation (FIRO) (Delaney *et al.*, 2020; Zarei *et al.*, 2021). Hydrologic predictions at different time scales are based on different processes (e.g., seasonal projection based on snow water storage, short-term prediction based on weather forecast) and subject to various levels of uncertainty. In addition, different forecast products have different lead-time (ranging from hours by short-term weather forecast to seasons by climate models), a better understanding of hydrometeorological factors at various time scales affecting reservoir operation would facilitate FIRO to select the forecast products suitable for a specific reservoir.

2.5 Conclusions

This chapter proposes a hierarchical temporal scale framework to improve the data-driven reservoir release modeling. When the dominant explanatory variables observed inflow or storage are unavailable as inputs, more than 60% of reservoirs across the CONUS gain the improvement in model performances, while modeling of 80% of them can be more accurate by this framework if the first layer is constructed at weekly scale. The proposed framework accounts for the influence of multiple temporal-scale variability to accurately

predict reservoir release behavior, which may have inspiring implications for data-driven reservoir release modeling in regions where operating records are incomplete or limited in availability.

This hierarchical framework is not model-specific and therefore has broad applicability. By further adjusting the primary states simulated on the first coarse scale, which is partially similar to the operating process of reservoir managers in response to the daily inflow corresponding to the predefined water control plans, the hierarchical architecture is conducive to improve both the performances and the interpretability of DDMs, and can be further adapted to be closely integrated with the decision-making of managers. It also demonstrates the similarity of a natural-human system and hydrologic processes across temporal scales. In future work, data-driven reservoir components that have comprehensive utilization of multi-timescale information could be incorporated into physics-based models to improve the accuracy of hydrological process simulations.

Results of different experiment settings reveal that reservoir inflow and storage volume have a paramount influence on the release strategies. Model-based sensitivity analysis proves it, and further illustrates that variable importance can vary in time periods and across multiple time scales. For nearly 1/3 reservoirs across the CONUS, reservoir operations mainly depend on different decision variables at different time scales, and for several reservoirs, such as some in the Upper Colorado, hydroclimatic forcing still has a major influence on the release, addressing more demands on the assessment and planning of reservoir status and accurate forecasting of hydroclimatic forcing.

Chapter 3

INCORPORATING SATELLITE-DERIVED VEGETATION INTERANNUAL VARIABILITY INTO LAND SURFACE MODELS

3.1 Introduction

Vegetation plays an important role in terrestrial water and energy budgets (Bonan, 2008; Duveiller *et al.*, 2018; Oehri *et al.*, 2022), global carbon cycle (Chen *et al.*, 2019b; Sha *et al.*, 2022) and adaptation to a changing climate (Duarte *et al.*, 2013; Thorne *et al.*, 2017; Anderson and Song, 2020). Vegetation transpiration directly affects terrestrial-atmospheric processes as the main component in latent heat flux, accounting for about 40-80% of latent heat flux (Coenders-Gerrits *et al.*, 2014; Wang *et al.*, 2014; Schlesinger and Jasechko, 2014; Good *et al.*, 2015). In addition, vegetation indirectly interacts with other hydrologic processes in response to climate forcing. For example, hydraulic redistribution through vegetation root systems regulates soil moisture to accommodate climate fluctuation (Amenu and Kumar, 2008; Yu and D'Odorico, 2014; Barron-Gafford *et al.*, 2017; Sun *et al.*, 2018). The vegetation shade and canopy snow interception change the land surface albedo and snow energy balance in cold regions (Link and Marks, 1999; Webster and Jonas, 2018; Malle *et al.*, 2019). These feedbacks subsequently propagate to catchment fluxes such as surface water yield, evapotranspiration and groundwater recharge (Vertessy *et al.*, 2001; Wattenbach *et al.*, 2 04). Furthermore, the detected alterations in vegetation reflect both natural factors (e.g., wildfire) and human activities (e.g., reforestation and deforestation) (Nobre *et al.*, 2016; Boulton *et al.*, 2022; Chen *et al.*, 2019a), which introduce further uncertainty when evaluating the effects of vegetation on the hydrological cycle. Therefore, improving the representation of vegetation dynamics and its interaction with climate forcings

and hydrologic processes are essential for predicting hydrologic fluxes and state variables in a changing environment (Miller *et al.*, 2006; Ruhge and Barlage, 2011; Kumar *et al.*, 2019).

Land surface models (LSMs) and hydrologic models typically present plant phenology through static multi-year average monthly climatology. For instance, the Unified Noah LSM (Chen *et al.*, 1996; Chen and Dudhia, 2001; Ek *et al.*, 2003), a stand-alone, uncoupled, 1-D column LSM used in the North American Land Data Assimilation System Phase 2 (NLDAS-2, Xia *et al.*, 2012), employs the climatological Green Vegetation Fraction (GVF, as the black solid line shown in Figure 3.1a) to parameterize vegetation dynamics using a composite 5-year (1985-1989) average derived from the Advanced Very High Resolution Radiometer (AVHRR) satellite (Gutman and Ignatov, 1998). Such climatology parameterization scheme only represents vegetation average seasonal variation and misses the interannual variability (as colored lines shown in Figure 3.1a), therefore failing to capture the actual vegetation responses to climate fluctuations and extremes (e.g., droughts) and leading to errors in estimates in the hydrologic fluxes and states (e.g., Tang *et al.*, 2012; Wattenbach *et al.*, 2 04).

In semi-arid and arid regions like the Upper Colorado River Basin (UCRB), the structure and distribution of vegetation can be highly regulated by precipitation patterns (Miranda *et al.*, 2011; Poulter *et al.*, 2014; Dannenberg *et al.*, 2019). Figure 3.1b illustrates the correlation between annual maximum GVF derived from the seven remotely sensed vegetation products and annual precipitation anomalies from the NLDAS-2 forcing. All seven remotely sensed GVF (details see Table 3.2) consistently show correlations with annual precipitation, indicating the vegetation response to climatic variability in the semi-arid UCRB. As global warming induces notable increases in atmospheric water demand and vegetation water consumption across diverse regions (Yuan *et al.*, 2019; Zhang *et al.*, 2020b), the association between precipitation and vegetation variability in drylands is expected to be enhanced (Zhao *et al.*, 2022). Furthermore, the mean sensitivity of

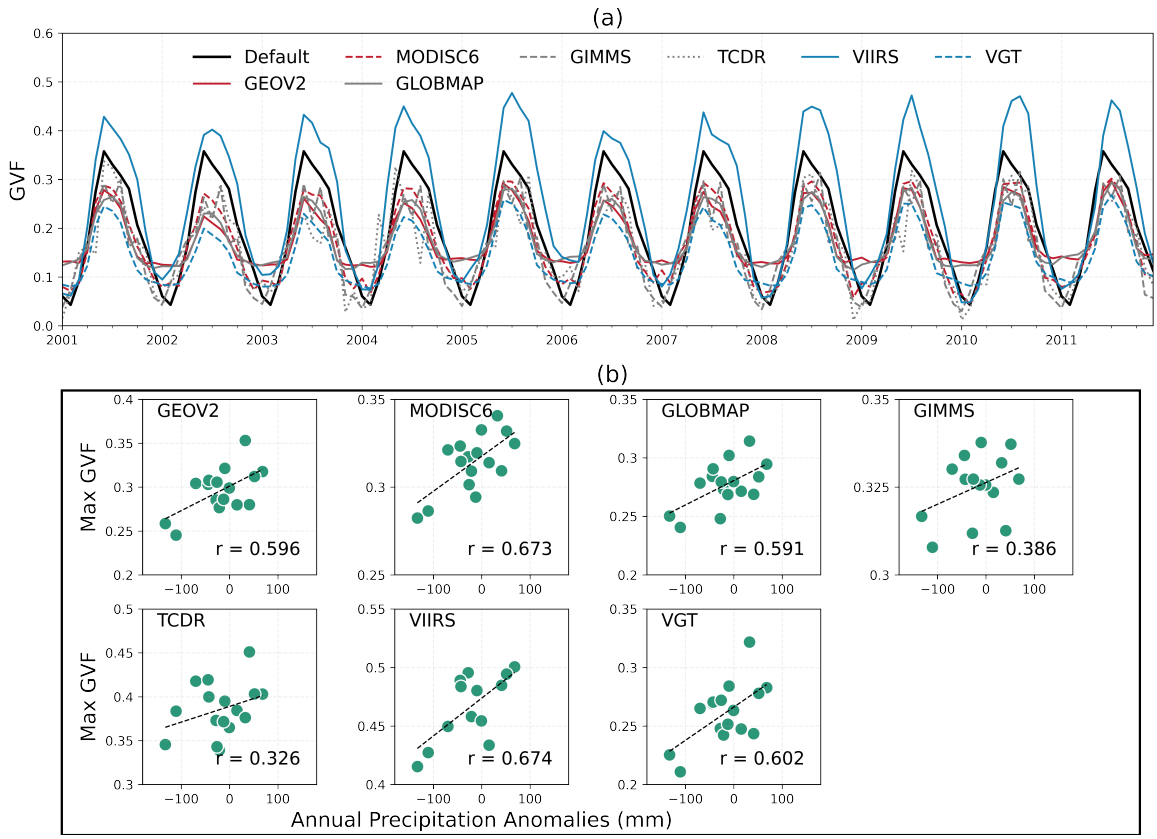

Figure 3.1: (a) Monthly average GVF during the period of 2001-2011 across the Upper Colorado River Basin (UCRB), here "Default" represents the climatological GVF obtained using a composite 5-year (1985-1989) GVF derived from the AVHRR; (b) Correlation between annual maximum GVF from various products (GEOV2, MODIS C6, GLOBMAP, GIMMS, TCDR, VIIRS and VGT) and annual precipitation. Details of remotely sensed products refer to Table 3.2.

vegetation canopy greenness to precipitation, as observed from satellites, is highest in arid regions and has been increasing over time in many drylands, including the western U.S., since the 1980s (Zhang *et al.*, 2022). The consistent observational evidence of the interaction between vegetation dynamics and precipitation suggests that the projected variations in precipitation could both directly adjust the water input to the system and alter vegetation indices, thereby introducing additional variance in hydrologic response. As a result, it is crucial to integrate interannually variant vegetation dynamics into LSMs for a more accurate representation of precipitation-vegetation association.

Mechanistic representations involve simulating vegetation dynamics through ecohydrological processes within the LSMs (e.g., dynamic vegetation module in the Noah-Multiparameterization Land Surface Model, Niu *et al.*, 2011). However, dynamic vegetation models are typically limited by simplified process representations and model parameterization uncertainty. Hence, the simulated dynamic phenology can be notably biased (Murray-Tortarolo *et al.*, 2013; Konings and Gentine, 2017) and often fails to reproduce the full extent of observed temporal variability (Koster *et al.*, 2014; Lee *et al.*, 2018; Kolassa *et al.*, 2020).

Remotely sensed time-varying vegetation indices are progressively incorporated into hydrologic models to yield a more accurate estimate of hydrologic states and fluxes. Enhancements in model performance, achieved by replacing the original climatological vegetation parameterization with real-time satellite datasets, are evident in simulated surface soil moisture and temperature (Yin *et al.*, 2016), streamflow predictions (Tesemma *et al.*, 2015; Ma *et al.*, 2019; Elmer *et al.*, 2022), and snow depth and terrestrial water storage (Kumar *et al.*, 2019). However, there are concerns associated with incorporating remotely sensed vegetation characteristics into large-scale hydrologic models. Discrepancies among various remotely sensed vegetation datasets can be substantial due to diverse sources of satellite sensors and retrieval algorithms. Although previous studies showed the benefits of fusing one specific vegetation index into hydrologic models, the propagation of uncertainty of multi-source remotely sensed vegetation products through hydrologic processes and impacts on hydrological simulations remains unclear (Cihlar *et al.*, 1997; Zhu *et al.*, 2018b). Moreover, the incorporation of multi-source remotely sensed data into modeling is a time-consuming process. There is a need for rapid diagnostic tools to aid in the evaluation and selection of appropriate remotely sensed data.

Machine Learning (ML) has gained considerable traction as a method to expedite hydrological simulations in recent years. It learns the relationships between various

independent and dependent variables. When the target outputs of the ML models are sourced from process-based models, the ML serves as a surrogate, effectively "replicating" the underlying mechanisms of the physical model. This surrogate role allows them to directly produce target outputs, achieving both high predictive accuracy and reduced computational runtime, as demonstrated in studies by Zahura *et al.* (2020); Zahura and Goodall (2022); Tran *et al.* (2021); Maxwell *et al.* (2021); Condon *et al.* (2021). Additionally, ML surrogates facilitate speeding up parameter estimation and uncertainty quantification (Xu *et al.*, 2017; Sun *et al.*, 2023), thereby improving the forecasting accuracy and predictive capabilities of hydrological models. Therefore, leveraging the potential of ML is a promising approach for efficiently integrating multi-source remotely sensed data into large-scale hydrologic models.

This chapter aims to examine the spatiotemporal impacts of fusing multi-source remotely sensed time-varying vegetation dynamics into LSMs on hydrologic simulations in the UCRB. With diverse natural conditions (e.g., climate, seasonality, topography and vegetation types) and significant societal impacts in a changing environment, the UCRB provides a wide spectrum to comprehensively assess the consequences of vegetation-hydrology interaction and land management. This chapter aims to address the following questions: (1) Can multi-source remotely sensed vegetation products capture the vegetation responses to climatic interannual variability? (2) In the UCRB, what are the hydrologic consequences of explicitly represented time-varying vegetation dynamics? (3) How does the response of vegetation to climatic fluctuation regulate hydrologic biotic and abiotic components, and what are the implications to water resources in the UCRB? This chapter develops a deep learning-based surrogate model based on the Noah LSM as a computationally efficient diagnostic tool to facilitate the parameter calibration and fusing of remote sensing products and process-based hydrologic models.

3.2 Methods

3.2.1 Study Area

As a prominent river in the southwestern U.S., the Colorado River spans seven U.S. states and Mexico, catering to the needs of 40 million people and providing irrigation water for 5.5 million acres of agricultural land (Cohen *et al.*, 2013). The natural flow of the Colorado River is heavily influenced by snowpack in the Rocky Mountain headwaters subbasins, accounting for over 70% of the river's annual streamflow (Li *et al.*, 2017). The UCRB, which generates approximately 90% flow of the entire Colorado River Basin (CRB) (Jacobs, 2011; Xiao *et al.*, 2018; McCabe and Wolock, 2020), is the primary focus of hydrologists. This snowmelt-dominated catchment extends from the headwaters in the Rockies of Colorado and Wyoming to Lee's Ferry in Northern Arizona, with elevations ranging from approximately 4,300 m to 900 m and covering an area of about 280,000 km^2 (Tran *et al.*, 2022). The climate of UCRB exhibits considerable variation, transitioning from alpine conditions in the north and east to arid/semi-arid conditions in the south and west (Ficklin *et al.*, 2013). The UCRB experiences an average annual precipitation of 400 mm, predominantly accumulating in mountainous regions (Hibbert, 1979), and ranging from over 1,000 mm in the east to less than 250 mm in the west (Ficklin *et al.*, 2013). Water balance fluctuations in the basin are primarily driven by ET, which constitutes over 85% of precipitation on average (Vano *et al.*, 2012). The basin features a predominance of shrubland and various forest types, together comprising approximately 60% and 25% of the total UCRB land cover, respectively. The UCRB encompasses the Green, Upper Colorado, Glen Canyon, and San Juan River subbasins (Figure 3.2).

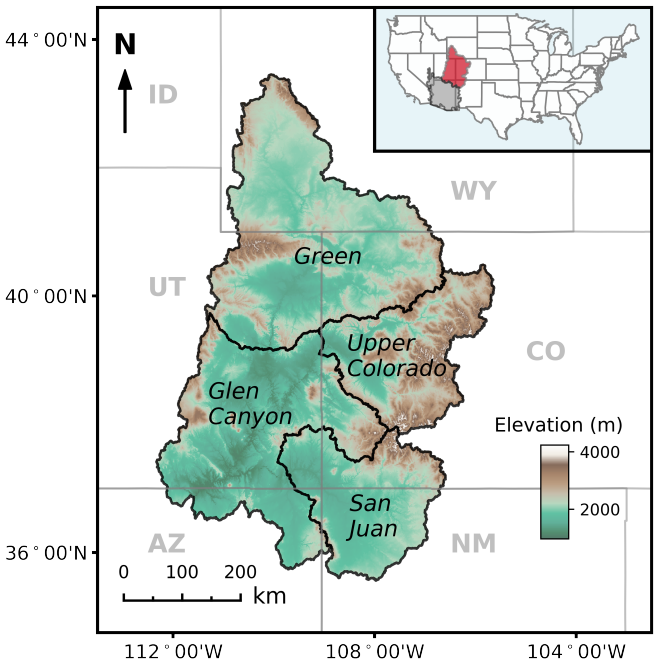

Figure 3.2: Location of Upper Colorado River Basin (UCRB) and subbasins.

3.2.2 Land Surface Model (LSM)

This chapter selects the NLDAS-2 Noah LSM version 2.8 as the focus for diagnostic analysis. The Noah LSM serves as one of the land surface modeling components within the NASA Land Information System (LIS,Kumar *et al.*, 2006), which typically operates in an uncoupled mode, utilizing a combination of observation-based precipitation, radiation, and meteorological and land surface parameter datasets. It is a one-dimensional soil-atmosphere-vegetation transfer model which simulates four-layer soil moisture (both liquid and frozen) with respective thicknesses of 0-0.1, 0.1-0.4, 0.4-1, and 1-2 m (Chen *et al.*, 1996; Chen and Dudhia, 2001; Ek *et al.*, 2003). Specifically, the Noah model version 2.8 is employed in the NLDAS-2 to generate long-term (1979-present), high resolution (1/8 degree) energy and water flux as well as state variables (Xia *et al.*, 2012; Kumar *et al.*, 2017). Although the operational version of NLDAS-2 is expected to undergo subsequent

enhancement via its transition to the NLDAS-Testbed at NCEP for future implementation by upgrading and/or replacing old versions and models and adding new assimilation procedures (Zhang *et al.*, 2020a), the diagnostic methodology presented herein remains readily transferable to alternative land surface models driven by NLDAS-2 forcing, and is amenable to applications within the context of other hydrological models.

In this chapter, particular emphasis is placed on the estimation of ET and its associated subcomponents, which are primarily governed by vegetation dynamics. Within the Noah LSM, the GVF dictates the partitioning of total ET into vegetation transpiration and evaporation originating from both bare soil and vegetative canopy. Additionally, GVF indirectly influences the snow sublimation rate by modulating state variables implicated in the water balance. In the following part of this section, the author briefly introduces the processes affected by vegetation in Noah LSM, while a comprehensive description of Noah LSM can be found in Chen and Dudhia (2001). The total ET consists of four components, the direct evaporation E_{dir}, the wet canopy evaporation E_c, the canopy evapotranspiration E_t and the snow sublimation, E_{sn}, namely,

$$ET = E_{dir} + E_c + E_t + E_{sn} \tag{3.1}$$

Specifically, the direct evaporation E_dir from the ground surface is computed by

$$E_{dir} = (1 - \sigma_f)\beta \cdot E_p \tag{3.2}$$

where E_p is the potential evaporation calculated by a Penman-based energy balance approach that includes a stability-dependent aerodynamic resistance (Mahrt and Ek, 1984), β is the first layer's soil moisture normalized by the field capacity and wilting point, and σ_f is the GVF.

The wet canopy evaporation E_c is determined by

$$E_c = \sigma_f \left(\frac{W_c}{S} \right)^n \cdot E_p \tag{3.3}$$

where W_c is the intercepted canopy water content, S is the maximum canopy capacity, and $n = 0.5$.

The plant transpiration E_t is determined by

$$E_t = \sigma_f B_c \left[1 - \left(\frac{W_c}{S}\right)^n\right] \cdot E_p \tag{3.4}$$

where B_c is a function of canopy resistance and is formulated as

$$B_c = \frac{1 + \frac{\Delta}{R_r}}{1 + R_c C_h + \frac{\Delta}{R_r}} \tag{3.5}$$

where C_h is the surface exchange coefficient for heat and moisture; Δ depends on the slope of the saturation specific humidity curve; R_r is a function of surface air temperature, surface pressure, and C_h; and R_c is the canopy resistance determined by

$$\begin{aligned} R_c &= \frac{R_{c\min}}{LAI \cdot F_1 F_2 F_3 F_4 \cdot s} \\ F_1 &= \frac{\frac{R_{c\min}}{R_{c\max}} + f}{1 + f} \quad \text{where} \quad f = 0.55 \frac{R_g}{R_{gl}} \frac{2}{LAI} \end{aligned} \tag{3.6}$$

where F_1, F_2, F_3, and F_4 represent the effects of solar radiation, vapor pressure deficit, air temperature, and soil moisture, respectively; s is the seasonal factor, determined by the optimum root growth temperature and mean soil temperature over the root zone; R_g is the incoming solar radiation, and R_{gl} is a lower limit of $30\text{W}\cdot\text{m}^{-2}$ for forests and $100\text{W}\cdot\text{m}^{-2}$ for crops (Noilhan and Planton, 1989); $R_{c\min}$ is the minimum stomatal resistance; $R_{c\max}$ is the cuticular resistance of the leaves. LAI is the leaf area index, and the calculation of seasonal LAI in Noah LSM is altered to depend on the vegetation class, rewritten as

$$LAI = LAI_{\min} + \alpha(LAI_{\max} - LAI_{\min}) \tag{3.7}$$

where $\alpha = \dfrac{\sigma_f - \sigma_{f\min}}{\sigma_{f\max} - \sigma_{f\min}}$ is normalized monthly GVF, σ_f is monthly GVF, and $\sigma_{f\max}$ and $\sigma_{f\min}$ are the maximum and minimum GVF, respectively. The details of vegetation

type-dependent maximum and minimum LAI and more detailed descriptions can be found in the work of Wei *et al.* (2012).

As the calculation of E_{sn} does not directly involve GVF, the details are not included here. For more comprehensive information, please refer to Livneh *et al.* (2010). To summarize, GVF affects ET fluxes in Noah LSM by determining the partitioning of latent heat between vegetation and soil, and further modulating vegetation transpiration by shaping canopy resistance.

3.2.3 Deep Learning-based Surrogate Model

This chapter develops a surrogate model to streamline the calibration of the Noah LSM, followed by a diagnostic analysis by fusing multi-source remotely sensed vegetation products into the calibrated surrogate. Surrogate modeling, or emulation, aims to provide a simplified and expedited model that reproduces the specific output of a complex model based on its inputs and parameters (Asher *et al.*, 2015).

The experimental design comprises two steps: process-based model running and surrogate model running (Figure 3.3). This chapter empirically identifies key parameters that influence the processes of energy partitioning and runoff generation. The minimum stomatal resistance ($R_{c\,\min}$) and the parameter used in the solar radiation term of the canopy resistance function (R_{gl}) (Eqn. 3.6) are also utilized as inputs. These parameters play crucial roles in determining canopy resistance, subsequently influencing the calculation of transpiration. Soil-related parameters fed into the model primarily include saturated hydraulic conductivity, saturated soil potential, porosity and the soil type "B" parameter. These factors are critical in computing the time rate of change of soil moisture and temperature. This chapter employs Latin Hypercube Sampling (LHS) to systematically select 1,000 parameter sets within the defined parameter space, following the NLDAS-2 specifications for soil and vegetation parameters (see Table 3.1). Wang *et al.* (2023) suggest

that the surrogate model is suitable for parameter optimization when the size of the parameter samples is approximately 40 times the number of parameters. LHS can ensure that the surrogate model encompasses a comprehensive range of possible parameter values. Subsequently, this chapter runs the Noah LSM using hourly NLDAS-2 meteorological forcing and all parameter sets to produce outputs of hydrological fluxes, including total runoff (Q) and four subcomponents of ET: transpiration (TRANS), canopy water evaporation (EVCW), bare soil evaporation (EVBS), and snow sublimation (SBSNO). These outputs aggregated from an hourly to a daily scale serve as target variables for the surrogate model in the subsequent step.

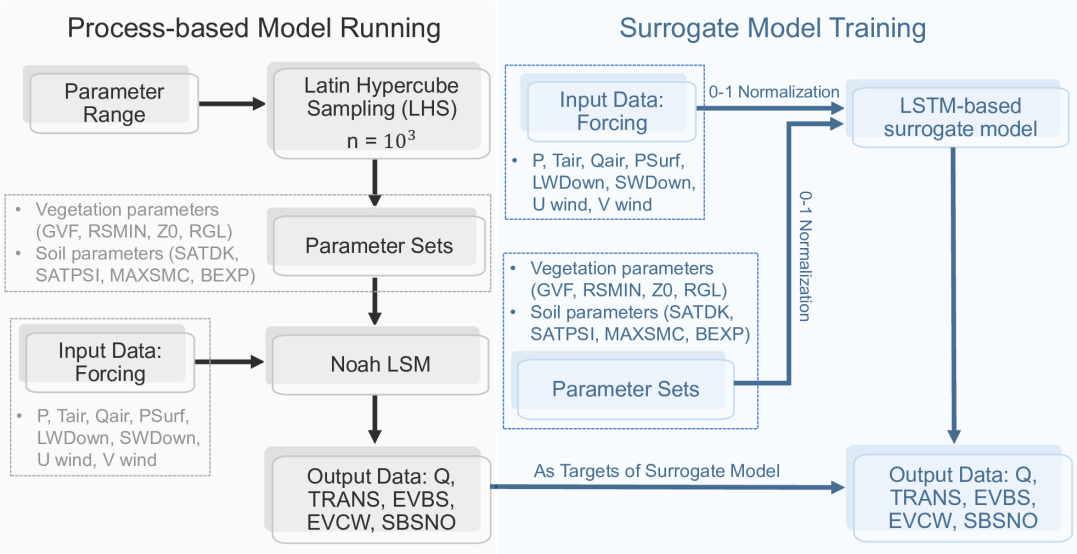

Figure 3.3: Flowchart summarizing the method for surrogate model construction. The meteorological forcing includes Accumulated hourly precipitation (P), Air Temperature (Tair), Specific humidity at 2 meters above the surface (Qair), Surface pressure (PSurf), Surface downward longwave radiation (LWDown), Surface downward shortwave radiation (SWDown), U wind component (U wind), V wind component (V wind). The physical parameters (vegetation and soil-related) include GVF, Minimum stomatal resistance (RSMIN), Roughness length (Z0), Parameter used in solar radiation term of canopy resistance function (RGL), Saturated hydraulic conductivity (SATDK), Saturated Soil Potential (SATPSI), Porosity (MAXSMC), Soil type "B" parameter (BEXP).

The surrogate model for the Noah LSM is established using Long Short-Term Memory (LSTM, Hochreiter and Schmidhuber, 1997), a deep learning technique adept at capturing temporal dependencies and extensively employed in hydrology in recent years (Kratzert *et al.*, 2018, 2019; Feng *et al.*, 2020; Ma *et al.*, 2021). LSTM is designed to capture both short- and long-term dependencies by the use of cell memory in its architecture and is frequently utilized to simulate hydrological fluxes (e.g., ET, Babaeian *et al.*, 2022; Jia *et al.*, 2023) and states (e.g., soil moisture, Orth, 2021; Li *et al.*, 2022; SWE, Meyal *et al.*, 2020; Song *et al.*, 2023). A comprehensive description of the LSTM architecture can be found in Section 2.2.3 Eqn. 2.2. This chapter builds a single-layer sequence-to-sequence LSTM model with 256 hidden units to simulate hydrologic fluxes. The sequence length encompasses the entire period, stretching from October 1st, 1979, to September 30th, 2016 during the training, validation, and test processes. This enables the surrogate model to mimic the underlying mechanism of a hydrologic model in a similar manner:

$$\begin{aligned} \text{Hydrologic Model:} \quad & y_t = f_{\text{Noah}}\left(x_t, s_{t-1}, \phi\right) \\ \text{Surrogate Model:} \quad & \hat{y}_t = f_{\text{surrogate}}\left(x_t, s_{t-1}, \phi; \theta\right) \end{aligned} \tag{3.8}$$

where x_t is the meteorological forcing at the time step t, s_{t-1} is the hydrologic or hidden states at the previous time step, ϕ is the physical parameters (soil and vegetation-related) in the Noah LSM, and θ represents weights and biases of the LSTM surrogate model, y_t is a subcomponent ET or runoff estimation of Noah LSM at the time step t, $\hat{y}_t$ is the subcomponent ET or runoff simulated by LSTM surrogate model at the time step t.

Similar to the process-based model, at each time step, the inputs for the surrogate model are meteorological forcing and parameter sets, while the outputs are hydrologic fluxes. Model inputs for the surrogate are normalized to a range of 0 to 1, to ensure faster convergence during the training process. Meteorological forcing data spanning from October 1st, 1979, to September 30th, 2016, is aggregated from an hourly to a daily scale. The fully connected operator enables the LSTM model to learn sufficient interactions between

meteorological forcing and physical parameters (e.g., GVF).

Table 3.1: Selected forcing variables and parameters used in the surrogate modeling.

Input Features	Feature Type	Unit	Range
Accumulated daily precipitation	Forcing	mm/day	-
Air Temperature	Forcing	K	-
Specific humidity at 2 meters above the surface	Forcing	kg/kg	-
Surface pressure	Forcing	Pa	-
Surface downward longwave radiation	Forcing	W/m^2	-
Surface downward shortwave radiation	Forcing	W/m^2	-
U wind component	Forcing	m/s	-
V wind component	Forcing	m/s	-
GVF	Vegetation	-	0-1
Minimum stomatal resistance (RSMIN)	Vegetation	s/m	0-400
Roughness length (Z0)	Vegetation	m	0-1.089
Parameter used in solar radiation term of canopy resistance function (RGL)	Vegetation	W/m^2	0-100
Saturated hydraulic conductivity (SATDK)	Soil	m/day	9.7394×10^{-7}-0.000176
Saturated Soil Potential (SATPSI)	Soil	m	0.035-0.7586
Porosity (MAXSMC)	Soil	-	0.395-0.476
Soil type "B" parameter (BEXP)	Soil	-	4.05-11.55

The objective function seeks to minimize the Mean Square Error (MSE) of all selected outputs (Q, TRANS, EVCW, EVBS, SBSNO) for both the LSTM surrogate model and the Noah LSM. Specifically, the objective function at each time step t is expressed as

$$\min_{\theta} \quad \alpha_i \sum_{i}^{5} (y_{i,t} - \hat{y}_{i,t})^2 \tag{3.9}$$

where i represents TRANS, EVCW, EVBS, SBSNO or Q, α_i represents the adjustable weights assigned to each of these outputs. α_i values allow for different levels of importance to be assigned to each output variable, which is crucial in multi-output models where some outputs might be more significant than others (e.g., the vegetation transpiration typically approaches zero during winter, whereas snow sublimation is generally low in the summer). The optimal weights and biases of the LSTM surrogate model are obtained through the training process using the Adam optimizer (Kingma and Ba, 2014).

The author opts for a space-split training approach, which involves using the entire time series of 50% of grid cells selected at random for training, 10% for validation, and the rest for testing. During the training process, initial hidden states in the LSTM surrogate model are set to zero, and the model is run on the first ten water years (1979-1989); the hidden states on September 30th, 1979, are employed as the initial conditions for subsequent training and fine-tuning. In the training of the model, the author employs a strategic learning rate schedule for optimization. For the first 100 epochs, the author sets the learning rate at 10^{-3} with a mini-batch size of 256 to facilitate rapid convergence. For the following 500 epochs, the learning rate is reduced to 10^{-4}, aiding in more precise adjustments in the model weights. According to the first 100 epochs' performance, the author adjusts the weights α_i in Eqn. 3.9 from a uniform set (where each weight is initially 1) to varied values through a trial-and-error method for optimal tuning. The author further decreases the learning rate to 10^{-5} to further fine-tune the model. During this fine-tuning phase, the author implements an early stopping mechanism. This approach monitors the model's performance on a validation set and halts the training process if there is no improvement, effectively mitigating the risk of overfitting. Once the model is well-trained, it is re-run from October 1st, 1979 to September 30th, 2016 to assess its performance during the test period (2000-2016). The performance of the model is evaluated using the Nash-Sutcliffe Efficiency (NSE, Nash and Sutcliffe, 1970), Root Mean Square Error (RMSE), and Percent Bias (PBIAS) (formulas provided in Eqn. 3.10). Results under five different random seed settings are evaluated for each training strategy to mitigate uncertainty arising from model implementation (random sampling and

weights initialization).

$$
\begin{aligned}
NSE &= 1 - \frac{\sum_{t=1}^{T} (y_t - \hat{y}_t)^2}{\sum_{t=1}^{T} (y_t - \bar{y}_t)^2} \\
RMSE &= \sqrt{\frac{\sum_{t=1}^{T} (y_t - \hat{y}_t)^2}{T}} \\
PBIAS &= \frac{1}{T} \sum_{t=1}^{T} \frac{y_t - \hat{y}_t}{y_t} \times 100
\end{aligned}
\tag{3.10}
$$

where y_t is the observed hydrologic flux and $\hat{y}_t$ represents the simulated hydrologic flux.

Utilizing the well-trained surrogate model, this chapter implements a Genetic Algorithm (GA) for effective parameter calibration. This chapter aims to determine an optimal set of soil-related physical parameters (specifically, SATDK, SATPSI, MAXSMC, and BEXP) to minimize the MSE between the simulated streamflow and the observed monthly streamflow data at Lee's Ferry, provided by the United States Bureau of Reclamation (USBR) naturalized flow records. To calculate the simulated monthly streamflow, the daily runoff outputs from the surrogate model are aggregated to a monthly scale and then summed across the entire basin, under the assumption that routing process impacts are negligible on this timescale. The GA systematically navigates through the parameter space (parameter range details in Table 3.1), in search of the optimal parameter set. In the GA configuration, the author specifies a population size of 300 with a maximum of 500 generations, a crossover probability of 0.8, and a mutation rate of 0.02. The selection process is based on roulette wheel selection.

The optimal parameter sets calibrated using the GA are then applied to the Noah LSM. Subsequent analysis is conducted utilizing this calibrated model. This chapter further validates the calibrated surrogate model by comparing its outputs with those of the process-based Noah LSM, both using the same set of calibrated soil-related parameters. If the surrogate model can reliably replicate the Noah LSM, the static GVF can be directly replaced with an interannually variant remotely sensed GVF dataset during data fusing in the calibrated surrogate.

3.2.4 Remotely Sensed Green Vegetation Fraction (GVF) Data Set

In order to evaluate the spatial and temporal impact of timing-varying vegetation dynamics on hydrologic simulations within the Noah LSM, several remotely sensed GVF datasets are fused into the well-trained surrogate model. This chapter employs both LAI-based and spectral-based satellite GVF products. Table 1 provides a summary of the data sources and resolutions for all seven products. LAI-based GVF datasets are derived from Système Probatoire d'Observation de la Terre VEGETATION (SPOT/VGT) Collection Version 2 (hereafter referred to as GEOV2, Verger *et al.*, 2013, 2014), Moderate Resolution Imaging Spectroradiometer Collection 6 (MODIS C6, Yan *et al.*, 2016), Long-term Global Mapping (GLOBMAP, Liu *et al.*, 2012), Global Inventory Monitoring and Modeling System LAI3g (GIMMS, Zhu *et al.*, 2013), and Terrestrial Climate Data Record (TCDR, Claverie *et al.*, 2016) LAI products. These datasets are converted using the following equation, in accordance with Norman's method (Norman *et al.*, 1995):

$$GVF = 1 - e^{b \times LAI} \tag{3.11}$$

where $b = 0.5$ is the extinction coefficient for general plant canopy.

The Visible Infrared Imager Radiometer Suite (VIIRS, Jiang *et al.*, 2016) and SPOT/VGT Collection Version 2 (herein VGT, Verger *et al.*, 2013, 2014) GVF datasets are derived from the presumed linear relationship between GVF and the Normalized Difference Vegetation Index (NDVI) or the Enhanced Vegetation Index (EVI).

All raw LAI or GVF tiles are post-processed including smoothing, gap-filling, and mosaicking to generate global maps at their original spatial resolution, which are subsequently aggregated to an average resolution of 1/8 degree. Monthly GVF fields are considered valid for the 15th day of each month. The study covers the common periods of all data products from the calendar years 2001 to 2011. The black solid line in Figure 3.1a represents the multi-year average monthly climatological GVF parameterization adopted by

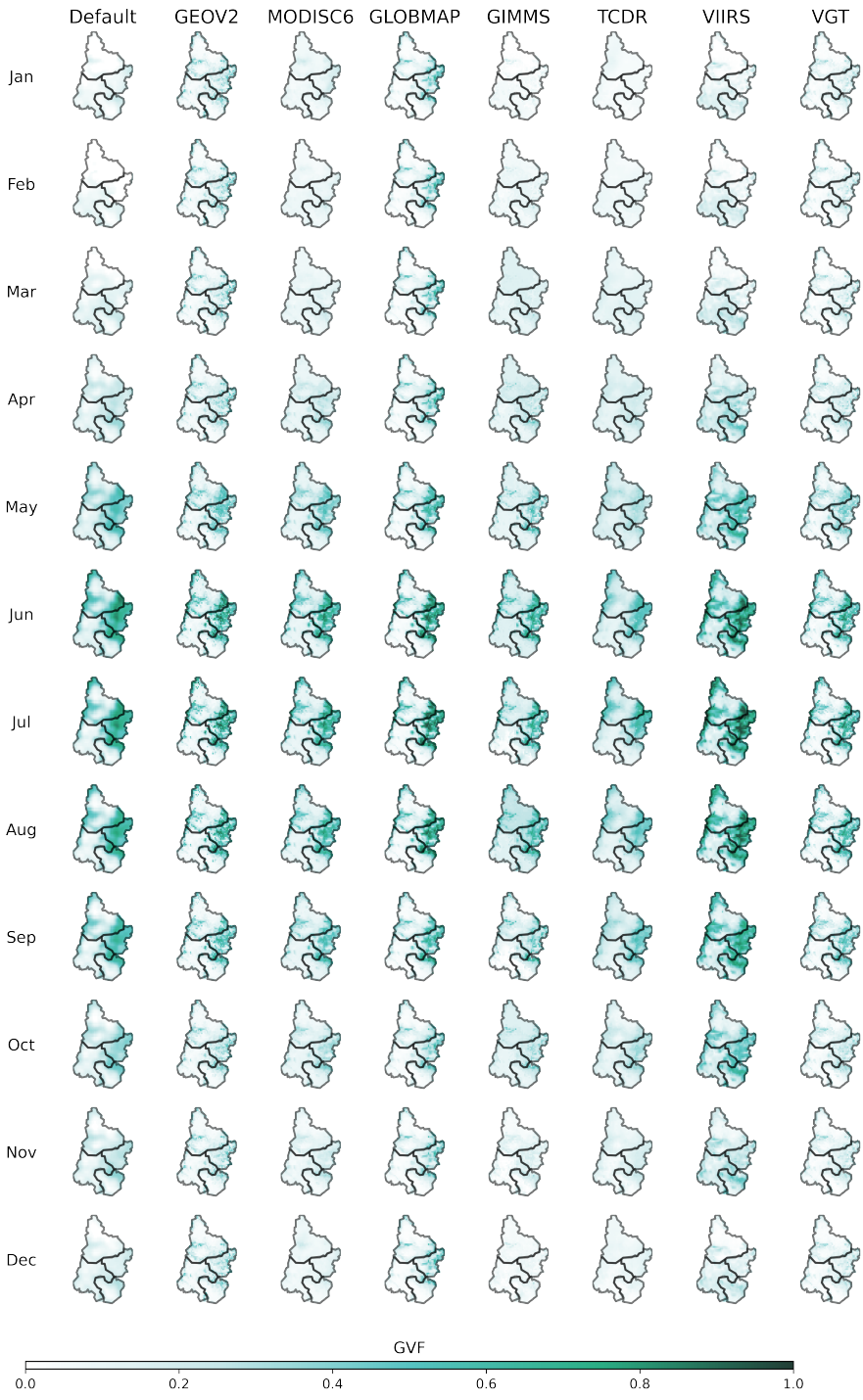

Figure 3.4: Spatial patterns of multi-year average (2001-2011) monthly remotely sensed GVF across the UCRB.

Table 3.2: Summary of remotely sensed GVF datasets employed in this chapter.

	Products	**Sensor**	**Spatial Resolution**	**Temporal Resolution**	**Period**
LAI-based	GEOV2	SPOT/VGT, MODIS	1 km	10 days	1999-
	MODIS C6	MODIS Terra	500 m	8 days	2000-
	GLOBMAP	AVHRR, MODIS	8 km	15 days (1981-2000); 8 days (2001-present)	1981-
	GIMMS	AVHRR, MODIS, SPOT/VGT	1/12 degree	15 days	1981-
	TCDR	AVHRR	0.05 degree	Daily	1981-
Spectral based	VIIRS	VIIRS	1 km	Daily	1981-2012
	VGT	SPOT/VGT, MODIS	1 km	10 days	1999-

NLDAS-2 Noah LSM, which is a composite of a 5-year (1985-1989) GVF dataset using the AVHRR satellite, based on the algorithm by Gutman and Ignatov (1998). Most of the processed interannually variant GVF products (colored lines, referred to as "dynamic GVF") are smaller than the default climatological GVF (referred to as "static GVF") during the growing seasons, but larger during the winter months (Figure 3.1a). Despite the differences in algorithms and original spatial and temporal resolutions, dynamic GVFs exhibit consistently similar seasonal patterns and temporal trends, except for VIIRS. This consistency is likely due to the fact that most vegetation indices are fused with other products. During the growing season, the domain average VIIRS GVF is noticeably higher than the static GVF and other dynamic GVF products. VIIRS channels have higher spectral and spatial resolution than AVHRR bands (Elmer *et al.*, 2022), potentially indicating a better representation of vegetation canopies and canopy gaps, partially due to improved

observation in high biomass regions with no apparent saturation in forested areas, reduced bidirectional reflectance distribution function effects, and decreased atmospheric influences (Ding and Zhu, 2018). In the following analyses, the author will also place particular emphasis on VIIRS, in addition to discussing the median statistics of the impact of all remote sensing products.

3.3 Results

3.3.1 Validation of Deep Learning-based Surrogate Model

The total runoff and ET subcomponents generated by the LSTM surrogate model show a great agreement with the Noah LSM simulations. The NSE values for total runoff (Q), total ET, TRANS, EVCW, EVBS, and SBSNO are 0.9914, 0.9943, 0.9971, 0.9973, 0.9963 and 0.9910, respectively. The RMSE values for all outputs remain below 0.15 mm/day, while the PBIAS values are within a range of +/- 2%. Here the total runoff (Q) and subcomponents of ET are directly obtained from the surrogate model output, while total ET is calculated as the sum of all the individual components. The LSTM emulation accuracy is deemed acceptable, exhibiting a promising capability to capture main hydrologic flux variables under the given meteorological forcing and hydrologic parameter sets.

This chapter then runs the process-based model and the LSTM surrogate using the calibrated soil-related parameters and compare their respective outputs. The calibrated model demonstrates good performance when compared to the USBR naturalized streamflow estimate in the Lee's Ferry (see Figure 3.5). It reproduces the overall seasonal trends and regional attributes of total runoff (Q) and ET components throughout the UCRB (Figure 3.5 and 3.7). The surrogate accurately emulates the bi-modal pattern of total ET, which mainly stems from that of bare soil evaporation. The simulation bias of snow sublimation may imply that the surrogate model has limited representations of snow processes in the

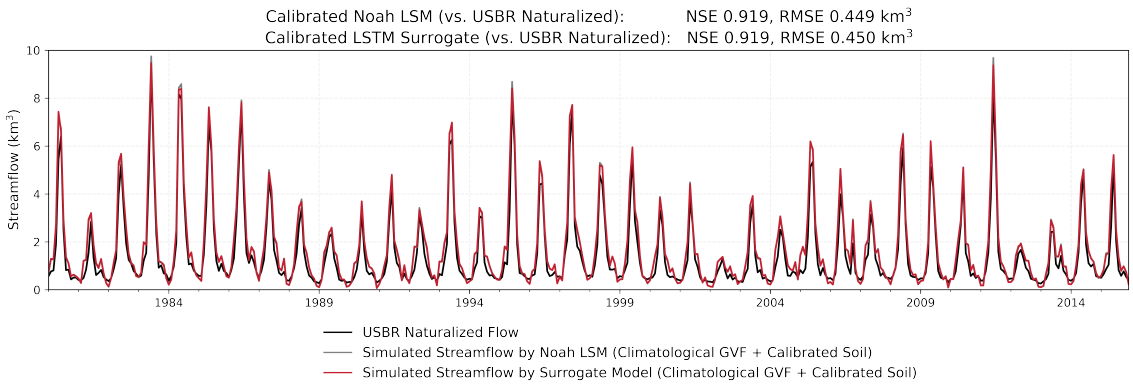

Figure 3.5: Calibrated simulated streamflow of Noah LSM and LSTM surrogate against the monthly naturalized streamflow for Lee's Ferry (1980-2015).

Noah LSM (Figure 3.6). This limitation could potentially impair the total runoff simulation of the surrogate model, considering the substantial contribution of snow-covered regions to the entire basin's water yield and the complicated snow processes involved in its runoff generation. While the surrogate model effectively tracks the increasing runoff from winter to spring and the subsequent decline through summer and fall for all four subbasins, it tends to slightly underestimate Q (Figure 3.6 and Figure 3.5). Despite these minor discrepancies, the LSTM surrogate model demonstrates a strong performance in simulating the temporal dynamics of ET and Q across the UCRB. After the integration of the remotely sensed GVF, the calibrated surrogate model continues to exhibit a high level of agreement with the outputs from the calibrated Noah LSM at the daily scale (Figure 3.8). This indicates that the surrogate model maintains reasonable extrapolation accuracy, even with the introduction of dynamic GVF data. Therefore, the subsequent analysis conducted using the LSTM surrogate model can be considered reliable.

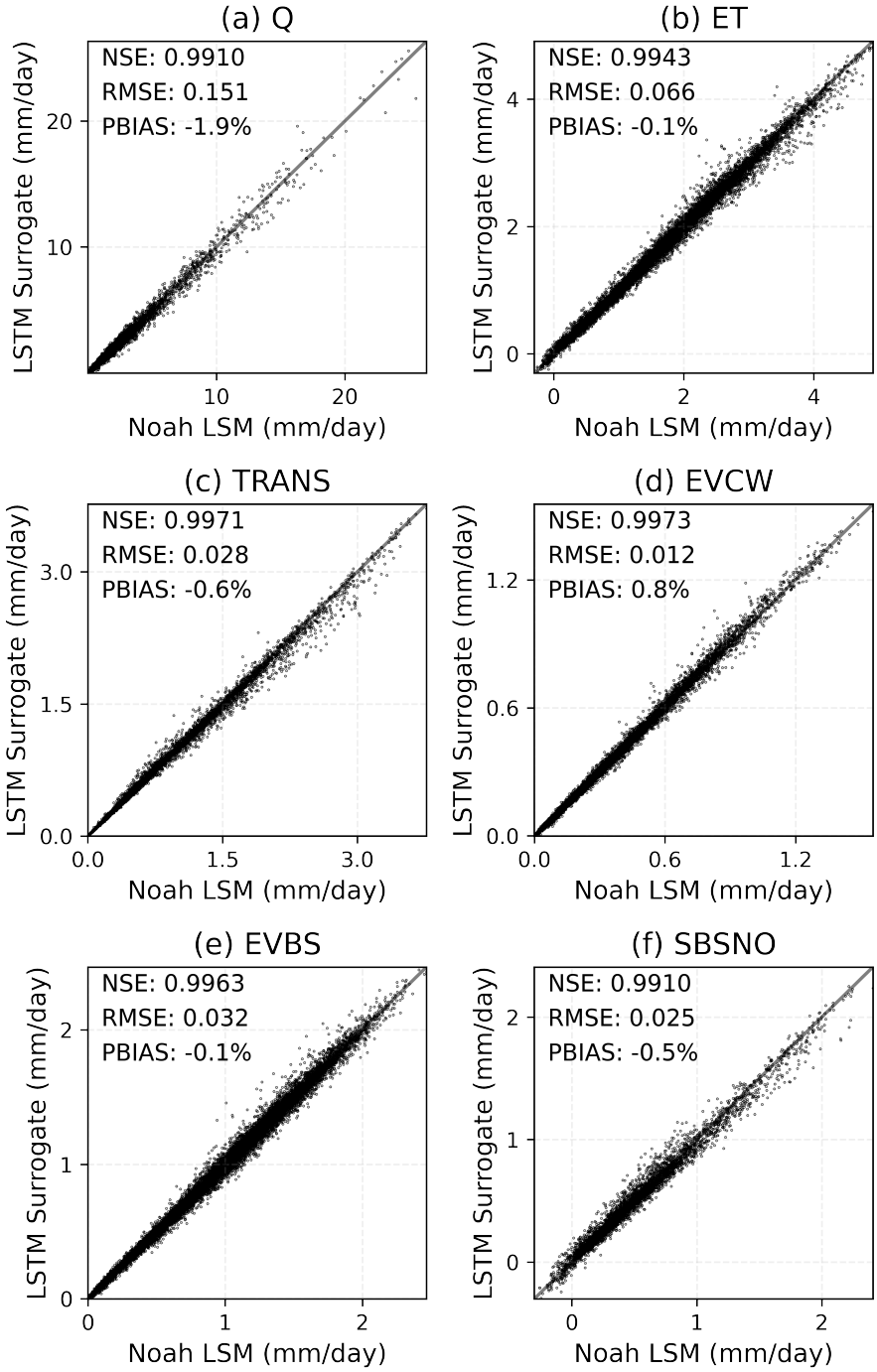

Figure 3.6: Performances of LSTM surrogate model for all selected target outputs, including (a) total runoff (Q), (b) total ET, (c) transpiration (TRANS), (d) canopy water evaporation (EVCW), (e) bare soil evaporation (EVBS), and (f) snow sublimation (SBSNO).

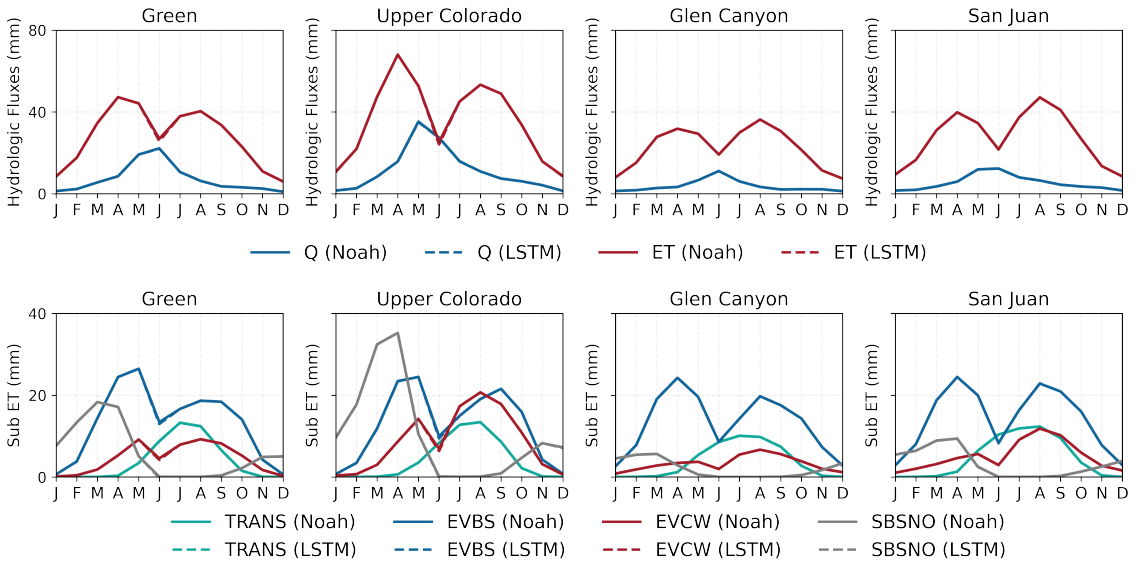

Figure 3.7: Comparison of multi-year monthly domain average (WY1979-2016) simulations of Noah LSM and LSTM surrogate model for all selected target outputs, total runoff (Q), total ET, transpiration (TRANS), canopy water evaporation (EVCW), bare soil evaporation (EVBS), and snow sublimation (SBSNO) at Green, Upper Colorado, Glen Canyon, and San Juan.

3.3.2 Effects of Incorporating Remotely Sensed Vegetation Dynamics into LSM

Figure 3.9 presents a comparison of differences in total runoff (Q), total ET, and its four components (TRANS, EVBS, EVCW, and SBSNO) before and after incorporating year-to-year variable remotely sensed GVF datasets in the UCRB during the 2001-2011 period. While remotely sensed GVF products differ in the interannual variability (Figure 3.12), most products (except VIIRS) consistently yield higher mean annual total runoff, ranging from +2.46 mm to +7.30 mm compared to Noah LSM with static vegetation parameterization (i.e., default climatological GVF). Although the differences in total Q magnitude are less than 10 mm basin-wise, it accounts for around 14% change relative to the multi-year average total Q of 56 mm in the UCRB during 2001-2011. The VIIRS results in a small decrease in total Q estimation (-3.81 mm). The two processed near-real-time GVF datasets from the same source, GEOV2 and VGT, exhibit a similar degree of increase in total

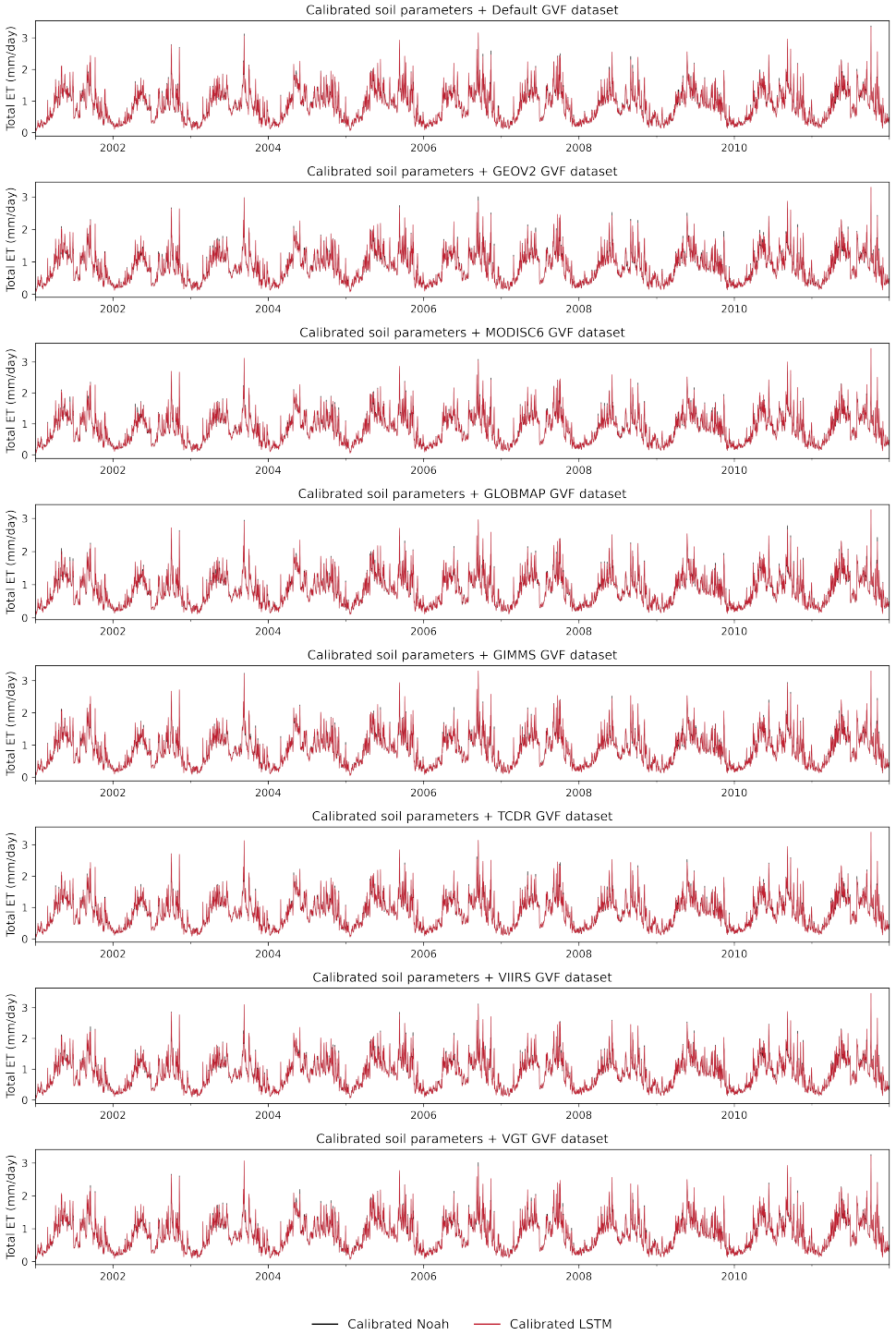

Figure 3.8: Comparison of daily total ET (2001-2011) generated by calibrated Noah LSM and calibrated LSTM surrogate model on a randomly selected grid cell.

Q after data fusion for 2001-2011, despite the former being LAI-based and the latter being spectral-based. The differences in total ET estimation after data fusion between GEOV2 and VGT may reflect the influence of GVF derivation methods (biomass or spectral) on hydrologic simulations in the UCRB.

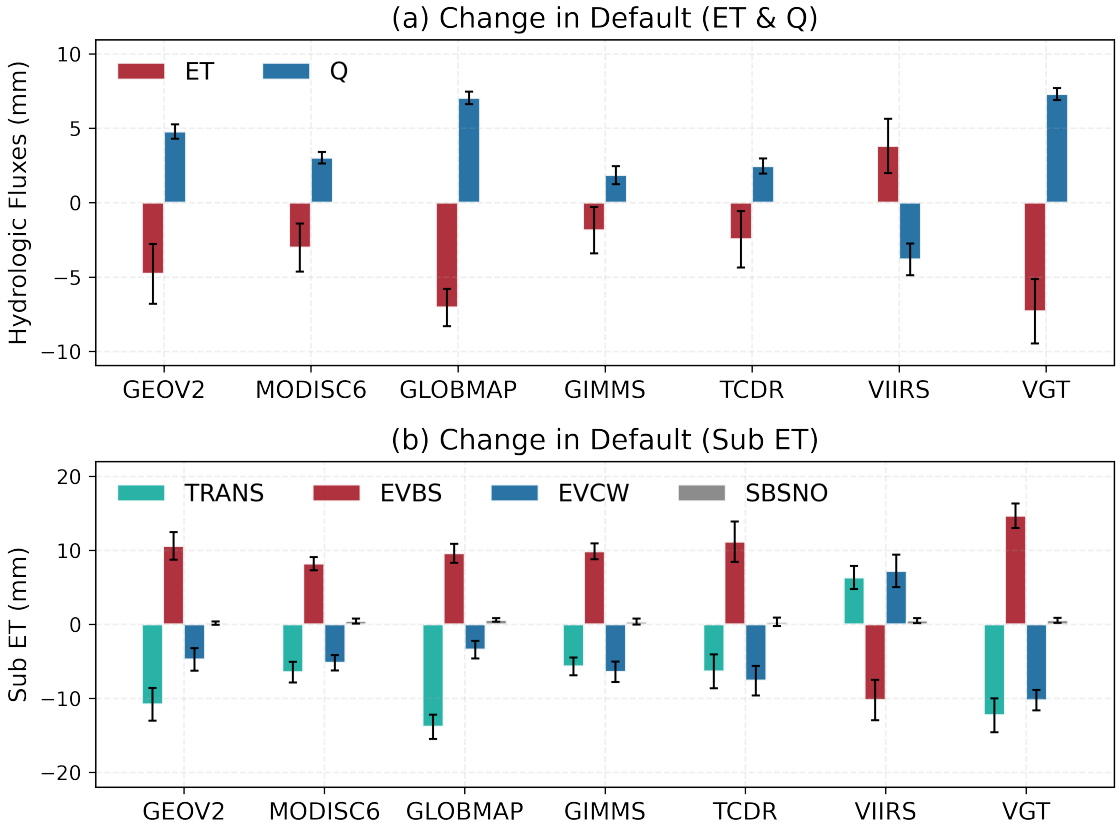

Figure 3.9: (a) Changes in mean annual ET and runoff (Q) in the UCRB after the integration of satellite-derived interannually varying GVF Products into the Noah LSM compared to that with static parameterization (i.e., default climatological GVF); (b) Analogous to Figure 3.9a, but for subcomponents of ET, including transpiration (TRANS), bare soil evaporation (EVBS), canopy water evaporation (EVCW), and snow sublimation (SBSNO). The color-coded bars represent the multi-year average values, while the black error bars indicate the standard deviations for the period 2001-2011.

For all satellite products, changes in total ET leads to nearly equivalent and opposite changes in Q in Figure 3.9a. This further indicates the LSTM surrogate model conserves the water balance at a long-term scale. Although other hydrologic variables are not simulated

by the surrogate model, it is expected that ET changes caused by dynamic vegetation characteristics will lead to changes in other processes. For example, the sensible heat or land surface temperature during the growing seasons is expected to decrease based on energy balance.

Figure 3.9b further shows the differences in various ET components caused by dynamic GVF parameterization. The satellite-based GVF consistently shows large bare soil evaporation than default monthly climatological parameterization (varying from 7 mm to 15 mm; except for VIIRS), which is partially compensated by decreased transpiration and canopy water evaporation. It may be attributed to lower GVF values from remote sensing during growing seasons than Noah's static parameterization. LSM model configurations (Eqn. 2) indicate that this smaller GVF leads to larger bare soil exposure, resulting in a substantial increase in bare soil evaporation among the four ET subcomponents. Upon integrating VIIRS, which presents higher GVF values during the growing season, bare soil evaporation experiences a decline. Nevertheless, the concurrent rise in transpiration and canopy water evaporation compensates for this decrease, leading to a minor increase in total ET (Figure 3.9a). Alterations in snow sublimation after fusing remote sensing products with interannual GVF variability are negligible. It may stem from GVF's indirect impact on snow sublimation within the Noah LSM.

Alterations in water budget terms induced by dynamic vegetation parameterization show inconsistent spatial patterns among different satellite products (Figure 3.10). In high-altitude regions in the Upper Colorado featuring woodland, wooded grassland, and evergreen needleleaf forest, the incorporation of remotely sensed GVF commonly triggers a positive change in annual total ET compared with Noah LSM employing monthly climatological GVF (except VIIRS). In the case of GEOV2, GLOBMAP, and VGT, the decrease in ET largely concentrates in shrubland regions within the southern UCRB. VIIRS GVF in the southern subbasins manifests an early onset and extended duration of greenness (Figure

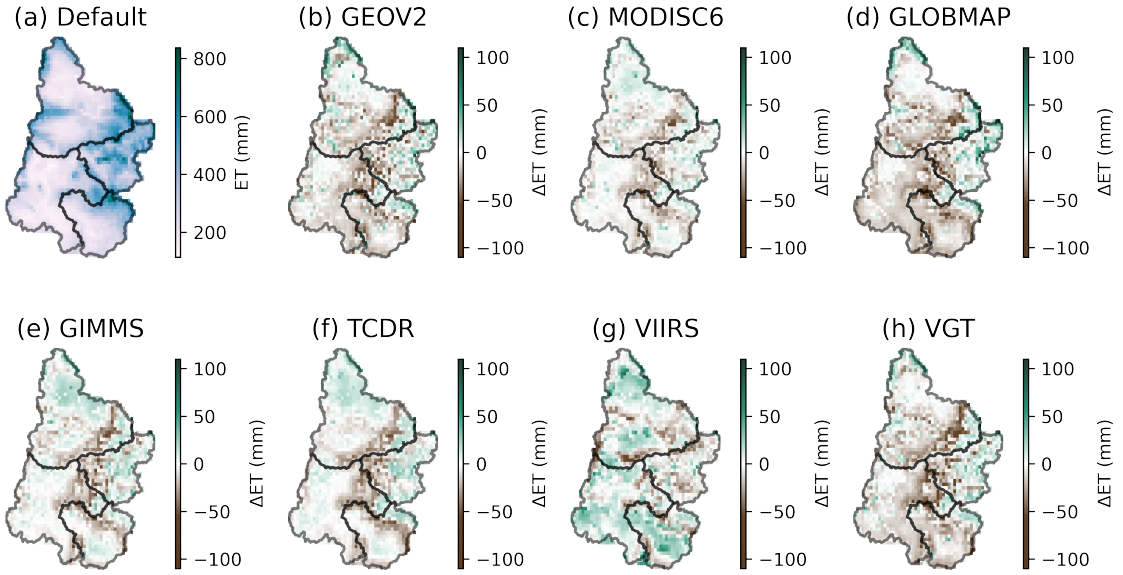

Figure 3.10: Spatial patterns of multi-year average (2001-2011) annual total ET across the UCRB for a) Default, b) GEOV2 – Default, c) MODISC6 – Default, d) GLOBMAP – Default, e) GIMMS – Default, f) TCDR – Default, g) VIIRS – Default, h) VGT - Default.

3.4), notably in San Juan. This leads to an elevated total ET estimate and a corresponding decrease in total runoff (Figure 3.11). Such spatial variations emphasize the diverse response of remote sensing products to vegetation types, a discrepancy that can propagate into hydrological flux estimates.

The impacts of dynamic vegetation on the interannual variability, indicated by the coefficient of variation (CV), of the annual total ET, vegetation transpiration, bare soil evaporation, and snow sublimation across the UCRB is shown in Figure 3.12 (GEOV2 as an example). While the integration of various remotely sensed GVF datasets into Noah LSM leads to slight changes in the water budget terms averaged across the entire basin, they substantially influence the interannual variability of hydrologic fluxes over space. For GEOV2, in most regions except the northern Green and high-altitude regions in the Upper Colorado, satellite-derived dynamic vegetation characteristics increase the interannual variability of total ET and reduce that of total Q. The increased interannual variability

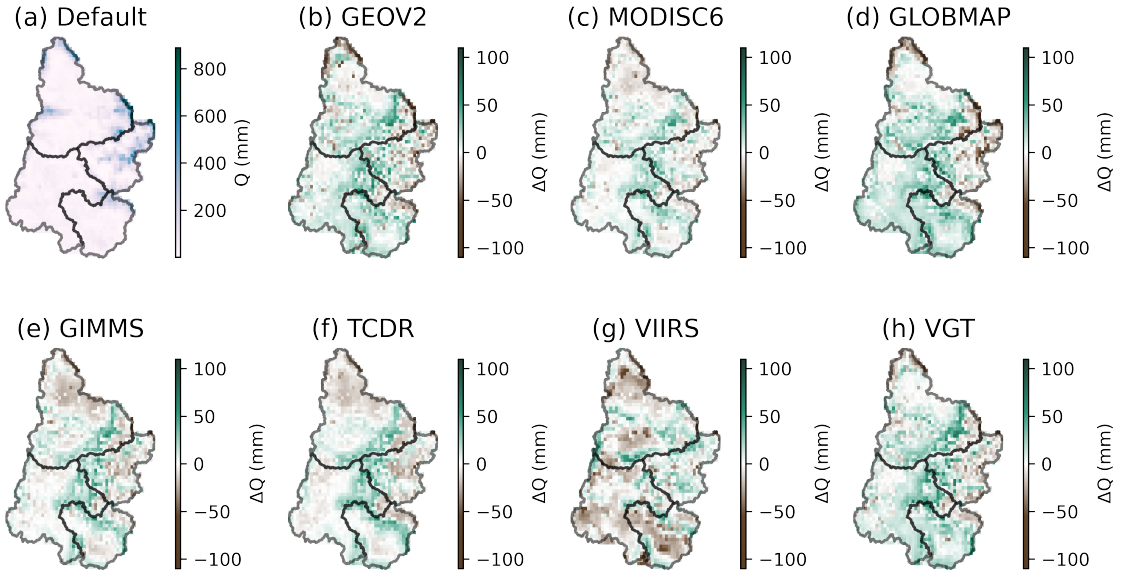

Figure 3.11: Spatial patterns of multi-year average (2001-2011) annual total runoff (Q) across the UCRB for a) default, b) GEOV2 – default, c) MODISC6 – default, d) GLOBMAP – default, e) GIMMS – default, f) TCDR – default, g) VIIRS – default, h) VGT - default.

in total ET mainly comes from the biotic subcomponents, with increased variability in vegetation transpiration and canopy water evaporation likely mirroring the actual response of vegetation (as captured in remotely sensed GVF products) to climate fluctuations. The reduction in interannual variability of bare soil evaporation across most regions indicates that the surface soil moisture can be more stable year by year. It can be attributed to dynamic vegetation characteristics that buffer (or exhibit resilience against) fluctuations in hydrologic fluxes and states. This highlights the ability of remotely sensed GVF products in capturing physiologic buffering effects against climate variability. The role of vegetation as a climate buffer highlights the importance of updating time-varying vegetation information in LSM simulations to better represent the hydrological variability.

Differences in the seasonal patterns of hydrologic fluxes in four sub-basins of UCRB between the Noah LSM using monthly climatological GVF and dynamic parameterization are primarily in the growing season (April to September), as illustrated in Figure 3.13a and

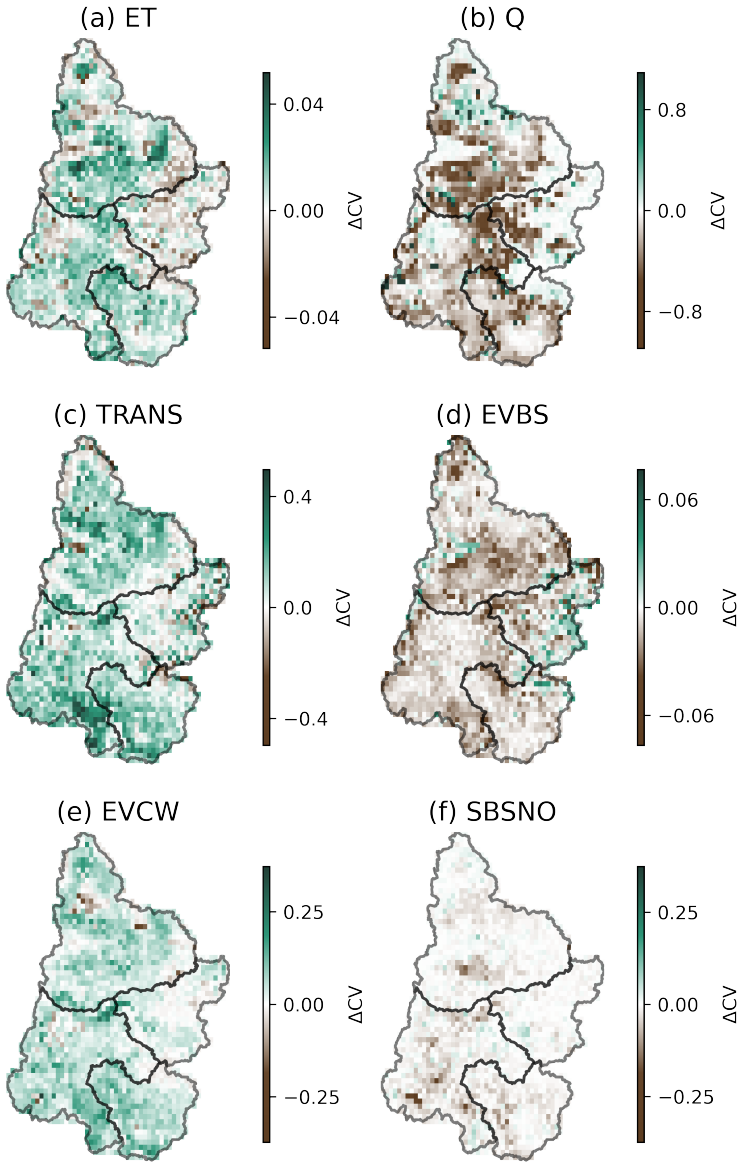

Figure 3.12: Spatial patterns of changes in Coefficient of Variation (CV) of (a) annual total ET, (b) annual total runoff (Q), (c) transpiration (TRANS), (d) bare soil evaporation (EVBS), (e) canopy water evaporation (EVCW), (f) snow sublimation (SBSNO) across the UCRB for GEOV2 product compared to default climatologic GVF parameterization.

S5. By incorporating satellite vegetation dynamics, the total ET and Q of Upper Colorado undergo changes, resulting in a reduction in total ET of approximately -3 to -1 mm among different GVF products between July and August and an increase in runoff by around 3 mm during late summer (Figure 3.13a). The total ET in the Green, another subbasin situated in the northern UCRB, is predominantly influenced by time-varying GVF changes from June and August. Integrating interannually dynamic GVF leads to a modest decrease in total ET for the two southern subbasins (i.e., Glen Canyon and San Juan), characterized by a bimodal pattern with peaks in April and August, while total runoff correspondingly increases.

An important seasonal feature of the UCRB is the bi-peak of bare soil evaporation and single peak of transpiration, as shown in Figure 3.13b. The soil evaporation reaches the first peak in April or May with high soil moisture content due to snow melt, followed by the vegetation transpiration peaks in June or July. The bare soil evaporation has the second peak at the end of growing season around August or September. Depending on the relative magnitude of bare soil evaporation and vegetation transpiration, the total ET has two peaks across all subbasins. In addition, bare soil evaporation is the largest among all ET components in the UCRB, and GVF directly changes the partitioning of total ET into bare soil evaporation and transpiration as indicated by Eqn. 3.2.

Interannually variant GVF does not markedly shift the seasonal phases of hydrologic variables, but does induce moderate modifications in their magnitudes, particularly for the ET partitioning. The transpiration peak consistently occurs in July or August, with amounts ranging from 10 mm to 13 mm over years, and the differences resulting from the integration of remotely sensed GVF vary between -2 mm and +3 mm among different products. Compared to the northern subbasins (i.e., Green and Upper Colorado in Figure 3.13a), vegetation transpiration experiences a more pronounced effect in the southern subbasins (i.e., Glen Canyon and San Juan in Figure 3.13b). Evaporation (both from bare soil and canopy surfaces) shows a larger alteration than transpiration in response to the

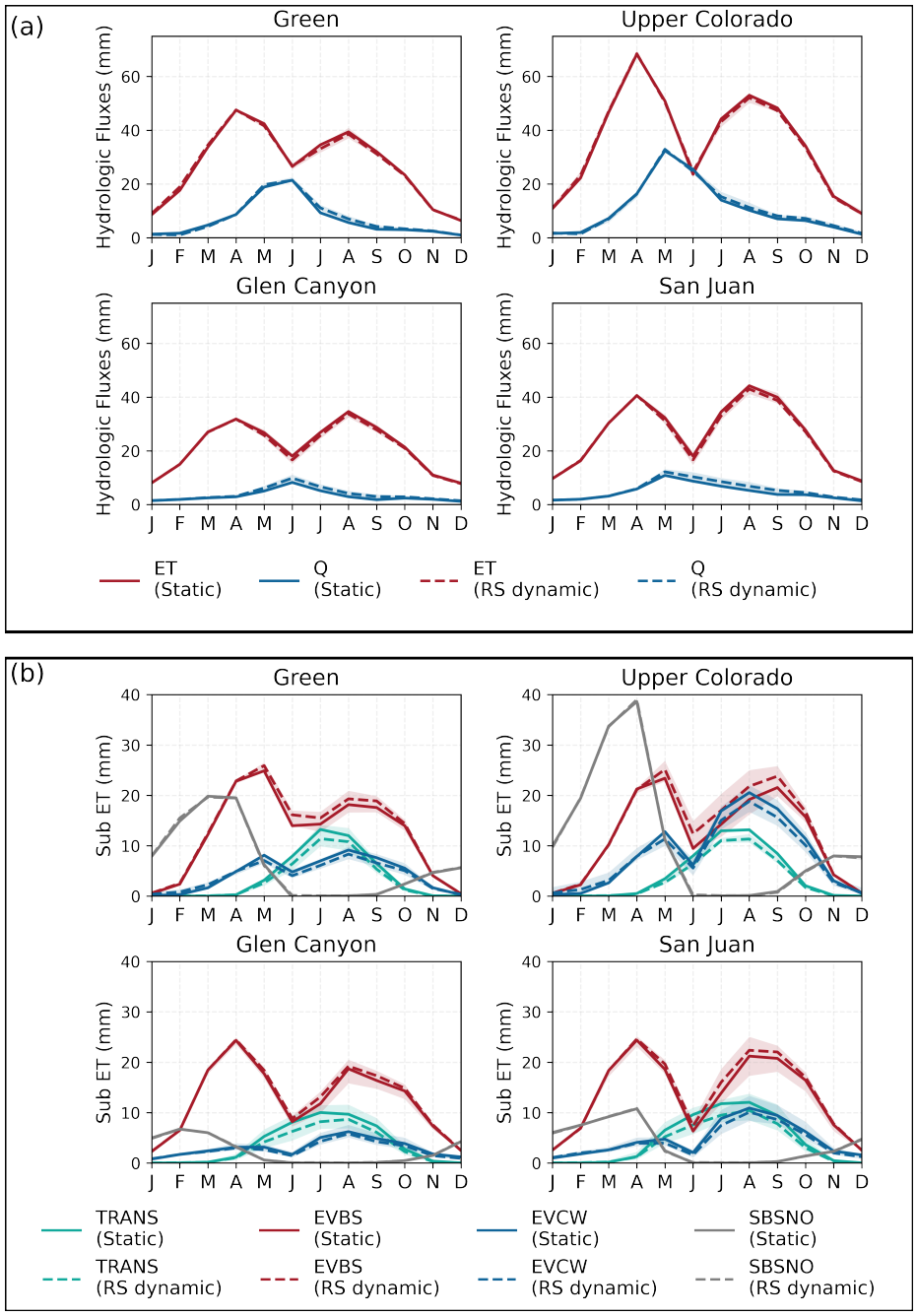

Figure 3.13: (a) Multi-year domain average total ET and total runoff (Q) across the four subbasins Green, Upper Colorado, Glen Canyon, and San Juan in the UCRB during 2001-2011. (b) Analogous to (a), but for subcomponents of ET, including transpiration (TRANS), bare soil evaporation (EVBS), canopy water evaporation (EVCW), and snow sublimation (SBSNO). Shading area depicts the uncertainties from the incorporation of seven remotely sensed GVF products, while the dashed line represents the median of these products.

year-to-year variable vegetation dynamics.

Uncertainties from remote sensing GVF products can propagate into total ET and runoff estimates, particularly evident in Upper Colorado during the late summer seasons (Figure 3.13a). These uncertainties become more prominent for ET subcomponent estimates, especially during growing seasons, as evident in Figure 3.13b, where the dashed lines, representing the median of the seven products, significantly diverge from the solid line (Noah LSM with static scheme). Such uncertainties, originating from various data sources, can be approximately interpreted as the sensitivity of hydrological flux variables to GVF changes. During non-growing seasons, these hydrologic variables display insensitivity to GVF changes. As shown in Figure 3.1 and 3.4, during winter months, most of the remotely sensed GVF values used in this chapter are higher than the climatological GVF. However, the associated response of ET subcomponents is marginally affected by these differences (almost negligible), possibly due to that vegetation growth activity, which plays a role in energy partitioning processes, is inactive during wintertime. In summary, from April to August, dynamic vegetation changes influence total ET by reallocating among ET subcomponents (mainly affecting bare soil evaporation), consequently impacting runoff in the UCRB and its four subbasins.

3.4 Discussion

3.4.1 Interactions between Parameters for LSMs after Updating Vegetation Information

While satellite products may more accurately reflect actual vegetation conditions, incorporating remotely sensed vegetation dynamics into the process-based Noah LSM does not automatically enhance the realism of the model. In this chapter, combined with the calibrated soil parameters, the performance of simulated streamflow even degrades after substituting the climatological GVF with an interannually varying GVF (Figure 3.14). This

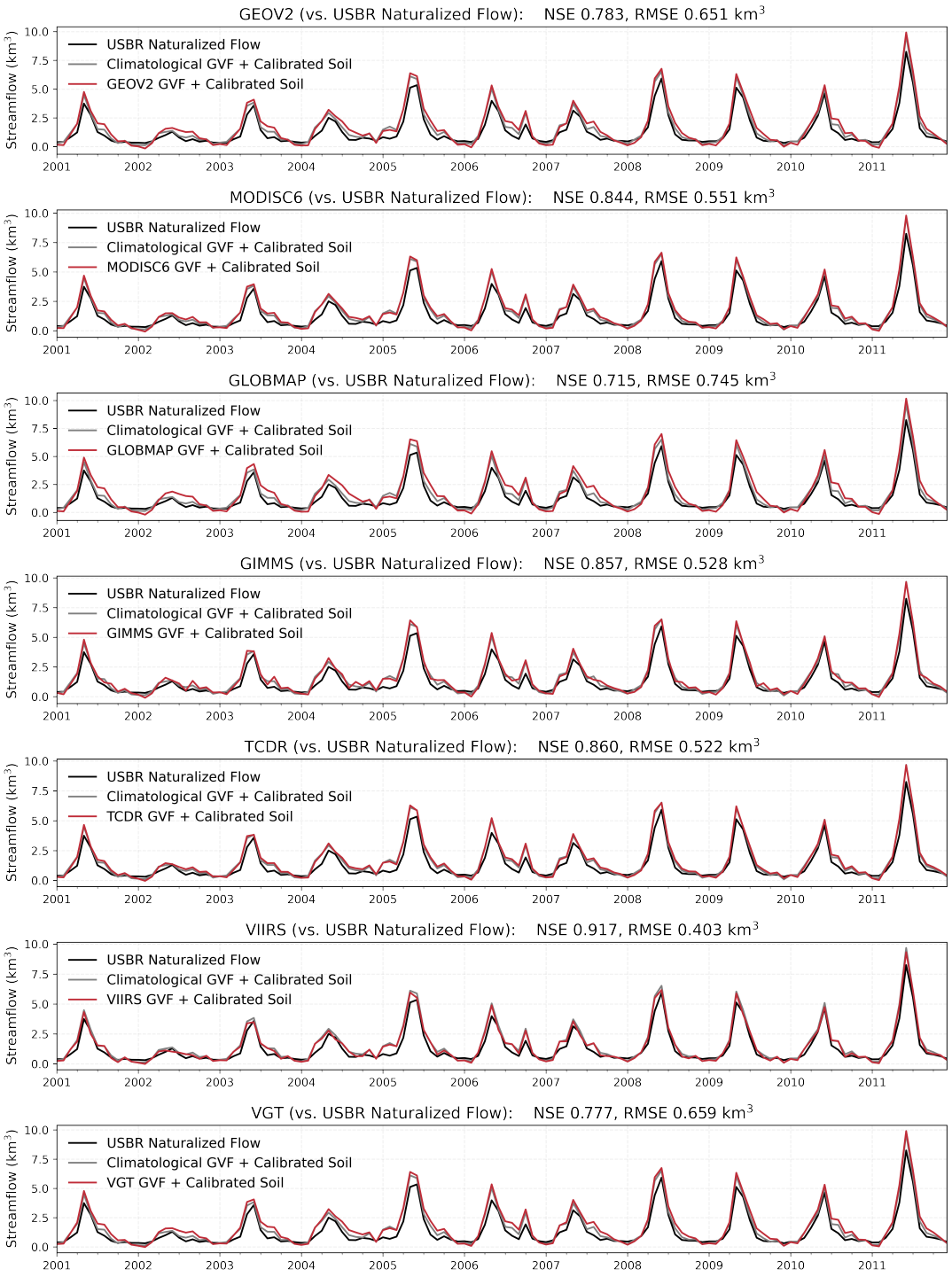

Figure 3.14: Comparison of simulated streamflow after the integration of remotely sensed GVF into Noah LSM against the monthly naturalized streamflow for Lee's Ferry (1980-2015).

can be attributed that data integration can modify the parameter-hydrologic component relationships within the model. These changes can manifest in two different ways. On one hand, vegetation in LSMs is a key determinant of the physical process how energy and water are distributed between biotic and abiotic components, thereby directly influencing hydrologic flux estimates (e.g., as Eqns. 3.2-3.4 shows in this chapter). In other hand, parameters can interact with each other within the model. For example, the presence of vegetation can influence soil physicochemical characteristics, such as hydraulic conductivity and soil porosity (Zhang *et al.*, 2021; Qiu *et al.*, 2022). Concurrently, vegetation tends to adapt to different soil types (George *et al.*, 2012; Rao *et al.*, 2016), indicating that the physical and chemical properties of the soil can also significantly shape and affect the type of vegetation that thrives. When vegetation information is changed from a climatologic to a dynamic parameterization, the interactions between vegetation and soil parameters correspondingly alter, therefore leading to alterations in hydrologic estimates. Hence, the updating vegetation-related parameterization using time-varying satellite products may require model parameter recalibration (e.g., saturated hydraulic conductivity, saturated soil potential, minimum stomatal resistance, and roughness length) and a re-evaluation of certain model assumptions (e.g., snow shading by high vegetation) (Nogueira *et al.*, 2021; Ruiz-Vásquez *et al.*, 2023).

Another concern arises from the uncertainty inherent in remotely sensed vegetation products. As discussed in Section 3.3.2, spatial variations in vegetation indices can be propagated to the subsequent estimates in hydrologic components and their interannual variability, leading to significant discrepancies among diverse satellite products. This presents a challenge in confidently updating LSM parameterization with dynamic vegetation data from a specific satellite product, as it is hard to guarantee the accuracy and reality of the selected one. For example, this chapter calibrates the soil parameter based on the climatologic GVF that may be considered "biased" compared to any yearly varied remotely

sensed GVF. It may also introduce bias into the calibrated soil parameter, thereby resulting in degraded performance in streamflow simulation. A feasible approach may be to employ multi-objective optimization in LSM calibration (Mostafaie *et al.*, 2018; Denager *et al.*, 2023). By extending the calibration targets beyond streamflow to include variables like Land Surface Temperature (LST) or Snow Water Equivalent (SWE), the calibration process could become more comprehensive.

3.4.2 Implication for Climate Change Assessment

As UCRB generates around 90% of runoff for the CRB, understanding the responses of vegetation to climate change and forest management is fundamental to the water security of the southwestern US. This chapter demonstrates the importance of dynamic GVF parameterization which alters the redistribution of precipitation among hydrologic fluxes, their interannual variabilities and seasonal patterns throughout the UCRB. Though remotely sensed products do not show vegetation long-term trends during last decades in UCRB, future climate conditions may introduce gradual or abrupt changes in vegetation parameterization.

Climate changes can introduce adaptive responses in vegetation that provide feedback to the climate system (Ramstein *et al.*, 2017). It implies that future vegetation phenology and type distribution may substantially differ from the historical patterns, e.g., earlier greening (Lian *et al.*, 2020), longer duration of the active vegetation-growing season (Kunkel *et al.*, 2004; Liu *et al.*, 2018; Grossiord *et al.*, 2022), shifts in vegetation types (Sturm *et al.*, 2001; Kelly and Goulden, 2008). Although some earth system models that provide climate change projections already incorporate dynamic models of vegetation or land cover (e.g., MIROC-ESM, HadGEM2-ES, IPSL-CM5B-LR), in offline LSMs or large hydrological models, the common practice has been to use LAI/GVF climatology derived from multi-year monthly averages over historical periods (Christensen *et al.*, 2004; Christensen

and Lettenmaier, 2007; Harding *et al.*, 2012; Currier *et al.*, 2023). Static configurations may not adequately capture the potential dynamic changes in vegetation over a long-term future projection, thereby challenging its predictability under climate change. Inadequate representations of vegetation dynamics in hydrological models and uncoupled LSMs potentially result in biased estimations of water availability.

It is worthwhile to consider replacing static vegetation climatology with time-varying vegetation characteristics in model configurations for more reliable climate change assessments. A feasible way involves establishing a statistically based empirical relationship using satellite-derived vegetation indices. Grounded in observational data, it allows the models to capture the evolving vegetation dynamics projected into the far future, although the extrapolation of past observations to the future does come with inherent uncertainty. Alternatively, some models, such as SWAT and Noah-MP, already come equipped with dynamic vegetation modules. These modules simulate vegetation dynamics based on a theoretical understanding of ecological and physiological processes. When appropriately calibrated, they can effectively reflect the shifting vegetation patterns and provide mechanistic insights into how vegetation might respond to a changing climate. It is of note that when implementing a dynamic vegetation configuration, there may be a need for concurrent updates to other information such as land use and land cover (LULC). As discussed in Section 3.4.1, the interplays between vegetation parameters and other model parameters (particularly those related to soil) are likely to shift, requiring recalibration efforts.

3.4.3 Deep Learning-based Surrogate Model for LSMs

In physically based LSM model computing, the runoff generation process in grid cells generally operates independently without lateral interaction and is therefore feasibly parallelizable. Likewise, deep learning calculations can be efficiently executed on GPUs in

batches, significantly reducing computational time costs. In this chapter, a single forward run for all UCRB grid cells (over 2,000 at a 1/8th degree spatial resolution) during the entire simulation period (1979-2016) on a daily scale with the LSTM emulator takes mere 4 seconds using a single NVIDIA Quadro RTX 5000 GPU. This approach accelerates the parameter calibration process in this study, highlighting the effectiveness of deep learning-based surrogate models in expediting uncertainty quantification for process-based, distributed hydrological models, which typically demand substantial computational resources.

For successful surrogate model applications, both speed and accuracy are essential, necessitating large amounts of training data to achieve the required accuracy in replicating simulation outputs when constructing machine learning emulators (Kasim *et al.*, 2021). The computational cost challenges shift toward building a high-fidelity emulator with a limited range of training data. On one hand, the effectiveness of surrogate models in capturing the relationship between hydrologic inputs and outputs hinges on selecting an appropriate sampling approach and a suitable sample size. While the author utilizes LHS to generate 1,000 parameter samples in this study, a smaller sample size, approximately 20 times the number of parameters, is often adequate for pinpointing sensitive parameters (Wang *et al.*, 2023). This can further facilitate the more streamlined construction of the surrogate within a given uncertainty quantification framework, especially when constructing surrogate models at regional to continental scales with high spatial heterogeneity. On the other hand, the machine learning model architecture is crucial. The right architecture provides suitable priors for a given problem. LSTM-type architectures are well-suited for emulating surface hydrologic process model physics due to their similar structures involving outputs and states, whereas convolutional architectures (CNN) are natural solvers for Partial Differential Equations (PDE) and suitable for groundwater modeling surrogates. Combining the strengths of LSTM and CNN presents a novel possibility for building

distributed hydrological model emulators (e.g., Yang *et al.*, 2019a; Tang *et al.*, 2021), which could capture the spatiotemporal dynamics between hydrologic responses and predictors, particularly in cases involving surface-groundwater interactions.

Surrogate models exhibiting high accuracy can adeptly mimic the physics of specific process-based models, thereby expediting parameter calibration. This advantage can also be harnessed to the model recalibration process after replacing the static climatological parameterization with the time-varying satellite dataset. Nevertheless, surrogate-based parameter calibration has limitations. Firstly, there is a possibility of bias in surrogate models, as they may not fully replicate the complex processes of the original models. If the surrogate does not exactly represent the interplay between parameters and hydrologic inputs, the calibrated parameter derived from a biased surrogate could also be biased. Secondly, surrogates cannot address the intrinsic limitations of process-based models. For example, surrogate-based calibration does not inherently resolve the equifinality (i.e., different parameter combinations produce similar results, Beven, 2006) dilemma. Nevertheless, even with further recalibration efforts, the use of different remotely sensed vegetation products in place of climatological parameterization could lead to vastly different recalibrated parameter sets, though they may yield very similar evaluation metrics. Additionally, site-specific calibration often results in inconsistent and non-contiguous parameters across geographically similar, adjacent areas (Yang *et al.*, 2019b). This issue also remains unresolved with surrogate-based calibration. In response, recent developments like the differentiable parameter learning (dPL) approach proposed by Tsai *et al.* (2021) offer a novel method. This method bypasses traditional calibration, instead utilizing deep learning platforms to expedite parameter estimation (the hydrologic models have been directly implemented in the platforms supporting automatic differentiation). It preserves the physical integrity of the hydrologic processes and derives more meaningful and coherent parameter sets based on regional information.

Although surrogate-based parameter calibration has its drawbacks, it still serves as an efficient tool for complex or computationally expensive models, such as Variable Infiltration Capacity (VIC, Liang *et al.*, 1994) or other distributed models (Tsai *et al.*, 2021). It is particularly relevant considering the effort required to translate original LSMs (typically written in Fortran or C++) into Python for implementation in deep learning platforms.

Besides parameter calibration, the highly accurate surrogate model can be effectively employed as a diagnostic tool for generating hydrologic predictions. This enables easy implementation of data fusion and assimilation. Such a model simplifies the process of substituting specific data chunks, which might be challenging to integrate into a process-based model without extensive code modifications. This flexibility can broaden the scope for exploring new scientific questions. A practical example is the web-based HydroGEN platform (Condon *et al.*, 2021). The platform employs ML emulators to generate user-tailored seasonal to annual hydrologic scenarios for both groundwater and surface water systems by leveraging observations and advanced physics-based hydrologic models.

3.5 Conclusions

This chapter quantifies the impacts of dynamically varying vegetation on the spatial and temporal patterns of hydrologic processes, using the UCRB as a test case. The machine learning-based surrogate model accurately reproduces the physical processes represented by Noah LSM and allows computationally efficient fusing of vegetation parameters (i.e., GVF in this chapter) from multiple remotely sensed products. Without considering year-to-year variable vegetation dynamics, the static monthly climatological GVF configuration of Noah LSM misses the vegetation response to hydroclimatic variability, leading to underestimated water yield (as much as 14%). In this semi-arid and arid UCRB, various remote sensing products consistently predict that changes in biotic ET components (i.e., transpiration

and canopy evaporation) outweigh the changes in abiotic ET components (i.e., bare soil evaporation). As the abiotic ET component exhibits two peaks in April and August and biotic ET components have a single peak in July, the static monthly climatological GVF in Noah LSM may not fully represent the seasonal patterns in the UCRB.

Through incorporating satellite-derived time-varying vegetation indices into Noah LSM, this chapter highlights the importance of representing the response of vegetation to climate forcing variability. This chapter highlights the additional two mechanisms of vegetation in modifying hydrologic fluxes by 1) buffering effect to the climate interannual variability and 2) modifying the hydrologic seasonal patterns with different biotic and abiotic ET components. Therefore, vegetation responses to climate should be explicitly (either statistically or mechanistically) represented to predict the future hydrologic changes in the UCRB.

Chapter 4

ASSESSING THE EFFECTIVENESS OF EXISTING RESERVOIR MANAGEMENT AND IDENTIFYING REALLOCATION OPPORTUNITIES UNDER CLIMATE CHANGE

4.1 Introduction

Climate change is reshaping the hydroclimatic landscape, altering precipitation patterns (Muller and O'Gorman, 2011), evapotranspiration characteristics (Milly and Dunne, 2016), runoff regimes (Döll and Schmied, 2012) and terrestrial water storage (Pokhrel *et al.*, 2021). These alterations can affect future water availability and demand (Schewe *et al.*, 2014; Konapala *et al.*, 2020; Caretta *et al.*, 2022), raising concerns about water management under a changing climate (Milly *et al.*, 2008a; Cosgrove and Loucks, 2015). Reservoirs, as one of the fundamental water infrastructures, are not immune to these changes. In the nation, more than 92,000 dams constitute the cornerstone of water resources management for various purposes, including flood mitigation, water supply, navigation, hydroelectricity generation, recreation, and environmental flow (National Inventory of Dams, nd). However, the reservoir storage capacity and companion operation rules derived from historical hydrologic records (usually with limited record length) may fall short of accounting for the full spectrum of probable hydrologic variability and change (Ho *et al.*, 2017). For example, the 1922 Colorado River Compact was based on an anticipated average annual flow of 16.4 million acre-feet (MAF) measured from the relatively wet two decades at the beginning of the 20th century (Castle *et al.*, 2014). However, subsequent tree ring studies have revealed lower long-term average annual flows for the Colorado River, ranging between 13.2 MAF (Hidalgo *et al.*, 2000) and 14.3 MAF (Woodhouse *et al.*, 2006). As a result, Lake Powell

only reached its full capacity once in 1983 since its commission. Current reservoir systems on the Colorado River suffer from the overestimated water availability, which is further exacerbated by the ongoing Millennium Drought. Given inadequate historical records and non-stationary future climate, an inquiry emerges from reservoir operation stakeholders: Are *de facto* reservoir storage and operation rules, the backbone of modern water infrastructure, adequate to fulfill designed objectives under future climate?

Man-made dams and reservoirs function as coupled nature-human systems (Liu *et al.*, 2007), where both system capacity and operation policy are the outcomes of natural (e.g., hydrologic variations) and anthropogenic drivers (e.g., operation targets, construction costs). Reservoir operation policies are mostly derived from optimality under a set of objective functions and constraints to tradeoff between the hydrologic variability and operation targets (Giuliani *et al.*, 2021). For example, the objective of maximizing water storage for supply often conflicts with the need for flood control. During flood seasons, reservoirs must reserve certain storage space to accommodate potential floodwaters, reducing the hydroelectricity generation and water supply reliability for later uses (Krzysztofowicz and Duckstein, 1979; Jain *et al.*, 1992). Conversely, in dry periods, maximizing storage for water supply can compromise the reservoir's ability to handle sudden, heavy rainfall, posing flood risks (Croley and Raja Rao, 1979; Ding *et al.*, 2015). However, one of the key factors usually missing in the analysis is the water footprint from reservoirs (i.e., represented by evaporative loss), especially in arid- and semi-arid regions. Mekonnen and Hoekstra (2012) estimated that the blue water footprint of a mere 8% of global hydroelectric capacity is equivalent to 10% of the water footprint for global crop production in 2000. The quantity of reservoir evaporative loss is equivalent to 20% of the global annual consumption of water use (Zhao *et al.*, 2022). The non-beneficial evaporative water loss is also an outcome of this coupled nature-human system, arising from interplays between hydrologic variability (e.g., seasonal inflow patterns and potential evapotranspiration), reservoir characteristics (e.g., reservoir

head-area-volume relationship), and human management decisions (e.g., release schedules, conservation strategies). Despite the significant amount of reservoir evaporative loss, it remains unclear whether hydroclimatology and operation policy jointly affect reservoir evaporative loss and further contribute to the re-operation of reservoirs. Knowing the dominant factors and their temporal patterns in controlling reservoir evaporative loss will enable quantifying the tradeoffs among reservoir performances and the water footprint.

The challenges from the changing climate and potential benefits of reservoir re-operation pose the urgent need for examining options for adaptative reservoir management. Expanding reservoir infrastructure capacity would be one solution, but it is expensive under various physical and financial constraints (Iglesias and Garrote, 2015). Another more viable avenue is to reallocate existing reservoir storage capacity among various project purposes (Carriere and Wurbs, 1988; Wurbs *et al.*, 1990). Reallocation is the redistribution of a reservoir's water storage across its pools to address significant, long-term operational changes or specific requirements. The United States Army Corps of Engineers (USACE) has implemented 161 reservoir pool reallocations at 56 reservoirs for water supply by raising the top of the designated conservation pool, thus prioritizing water conservation purposes over flood control during dry seasons (Doyle and Patterson, 2019). Reservoirs with a distinct seasonal operation pattern would most likely benefit from seasonal pool reallocation, such as those in Texas. For example, a seasonal-varying operation rule curve was planned and implemented to replace the original constant conservation pool since 1990 in the Wright Patman Reservoir, TX, allowing the conservation pool elevation to increase from an original 220.6 ft to a higher level of 227.5 ft between April 1 and June 1. Summertime water storage capacity surged from 122,639 to 310,428 acre-feet, providing a much larger room for water reservation to mitigate drought conditions. Therefore, reservoir reallocation provides a promising approach for adaptative management under future climate with existing reservoir infrastructure. However, it remains unclear how to identify the strategies (e.g., timing and amount) of reservoir

reallocation to enhance the overall system performance.

With the pressing needs discussed above, this chapter aims to investigate the challenges and opportunities of reservoir infrastructure and operation policy under a changing climate. Specifically, this chapter will address the following research questions. First, are current reservoir infrastructure and operation policy sufficient to mitigate hydrologic changes under future climate to achieve designed operation goals? This chapter will build data-driven models to represent the reservoir manager's operation decisions. The data-driven reservoir operation models will be driven by projected hydroclimatic forcings to evaluate the system performances (e.g., flood risk reduction and water supply reliability) under future climate scenarios. Second, how do hydroclimatic variation, reservoir characteristics, and operation policy jointly determine the reservoir evaporative loss? A better understanding of reservoir water loss under interactions between hydroclimatic and anthropogenic drivers can further indicate the opportunities to evaluate tradeoffs and synergies among operation targets. This chapter will incorporate reservoir evaporation processes into the data-driven reservoir model to represent the interactions between natural and operational components of reservoir dynamics. Third, can opportunities be identified for reservoir re-operation strategies (e.g., reallocation of storage pools) to enhance the system performance with the de facto infrastructure? This chapter will focus on the strategies of seasonal-varying (i.e., in terms of timing and amount) reservoir conservation pool reallocation, which is mostly viable for real-world reservoir adaptive management. This chapter applies the generic framework and methods to reservoir systems in Texas as a case study due to the complexity of the hydroclimatic regime and the significance of mitigating hydrologic extremes (e.g., floods and droughts). In Texas, surface water makes up almost two-thirds of the total existing water supply (8.9 million acre-feet per year) for municipal, manufacturing, steam-electric, and mining users (Texas Water Development Board, 2021). The operation target may switch from flood control to water conservation within a short period (e.g., weeks) due to

the quick shift from flood season to dry season. The arid- and semi-arid climate also exemplifies the importance of evaporative loss in reservoir operation.

4.2 Methods

4.2.1 Data-driven Reservoir Operation Models under Hydroclimatic and Anthropogenic Forcings

This chapter focuses on a set of 21 reservoirs located in eastern Texas, which are owned and operated by the USACE Tulsa District and Fort Worth District (details in Table 4.1). To construct the data-driven reservoir operation models that mimic the existing regulatory rules, this chapter utilizes the standardized database for historical daily reservoir levels and operations of USACE reservoirs developed by Patterson and Doyle (2018). These observed records include daily reservoir storage volume (acre-feet, af), inflow (cubic feet per second, cfs), and release (cubic feet per second, cfs) for each reservoir. Gaps in the dataset are filled using nearest neighbor interpolation. The output variables of our data-driven models include reservoir release, evaporative loss, and water storage. The model forcings include reservoir inflows, air temperature (as a proxy for potential evaporation Livneh *et al.*, 2015), and reservoir water level (as a system state variable). Specifically, the author includes the current inflow, the previous 7-day inflow history, and accumulated inflow volumes for the preceding 15, 30, 90, 180, and 365 days. Recent instantaneous inflows represent the short-term hydrologic variability to capture flood mitigation-related release decisions, while accumulated inflow volumes represent the trend of water availability for water supply operations. The author also includes static inputs including reservoir conservation capacity and dead pool capacity to represent the reservoir infrastructure constraints. All water-related variables are converted to million cubic meters (mcm). This chapter applies the Long Short-Term Memory (LSTM, Hochreiter and Schmidhuber, 1997) to produce three daily

outputs: storage, evaporation, and release. While some studies use reservoir water storage as an input in data-driven models (Yang *et al.*, 2016, 2021; Longyang and Zeng, 2023b), it is noted that reservoir water storage level (a state variable) needs to be explicitly represented in the reservoir model to capture reservoir performance under future climate. This chapter updates daily reservoir storage in the simulation loop (i.e., simulated storage in the previous day storage is fed as one of the inputs for the current day's reservoir operation decision), which can be expressed as

$$\begin{aligned} h_t, c_t &= \text{LSTMCell}(X_t) \\ S_t &= \text{FC}(h_t) \\ E_t &= \text{FC}(X_t, S_t) \\ R_t &= \text{FC}(X_t, S_t) \end{aligned} \tag{4.1}$$

where LSTMCell is the LSTM layer, FC is the fully connected layer, X_t represents the model inputs, S_t is reservoir storage, E_t is the reservoir evaporation, R_t is the reservoir release, h_t and c_t are hidden states and cell memory in the LSTM, respectively. A comprehensive description of the LSTM architecture can be found in Section 2.2.3 Eqn. 2.2. This approach avoids calculating storage based on reservoir water balance where some water budget items (e.g., diversion and leakage) are not available. The author further implements the physics-constrained loss function to prevent potential unreasonable extrapolation, i.e.,

$$\begin{aligned} L = &\frac{1}{T}\sum_{t=1}^{T}\left(R_t - \hat{R}_t\right)^2 + \frac{1}{T}\sum_{t=1}^{T}\left(S_t - \hat{S}_t\right)^2 + \frac{1}{T}\sum_{t=1}^{T}\left(E_t - \hat{E}_t\right)^2 \\ &+ \lambda_1 \cdot \max\left\{\hat{S}_t - \text{FP}, 0\right\} + \lambda_2 \cdot \max\left\{\text{DD} - \hat{S}_t, 0\right\} \end{aligned} \tag{4.2}$$

where $\hat{R}_t$, $\hat{S}_t$, and $\hat{E}_t$ are simulated reservoir release, storage, and evaporation, respectively, FP and DD are the top of flood control pool and the top of the deadpool, λ_1 and λ_2 are the penalty coefficients which controls the importance of the penalty terms. The two penalty terms are utilized to ensure the simulated storage remains within a reasonable range: specifically, it does not exceed the flood control pool's top nor fall below the dead pool's

top. 60% of data records are used during the training process, 10% of them for validation, and the rest for testing. The Nash-Sutcliffe Efficiency (NSE, Nash and Sutcliffe, 1970) of the three outputs time series (storage, evaporation, and release) is applied to evaluate model performance. The well-trained reservoir operation models represent the actual reservoir operation decisions implemented by reservoir managers under the hydroclimatic variability and infrastructure capacity constraints.

The future climate forcings are dynamically downscaled six GCMs outputs (i.e., ACCESS-CM2, BCC-CSM2-MR, CNRM-ESM2-1, MRI-ESM2-0, MPI-ESM1-2-HR, and NorESM2-MM, details in the Supplementary Materials Table 4.2) in CMIP6 under the Shared Socioeconomic Pathways (SSP) 585 emission scenarios (Rastogi *et al.*, 2022). The bias-corrected 3-hourly climate forcings are fed into the calibrated Variable Infiltration Capacity (VIC) model version 5 (Hamman *et al.*, 2018) with 1/24° (around 4 km) spatial resolution to simulate runoff (Kao *et al.*, 2022). The author aggregates runoff generated in the upstream watershed of reservoirs to generate reservoir inflow. Future air temperature is obtained from Rastogi *et al.* (2022). It is noted that the future climate from SSP 585 represents a high CO_2 emission scenario with no mitigation policies (O'Neill *et al.*, 2016). Therefore, this chapter primarily focuses on the effectiveness of current reservoir management strategies under the most extreme future scenario in this chapter. In the assessment under future scenarios, all reservoirs retain their conservation pool capacity the same as that in 2023. The author also includes sedimentation scenarios following the projected rating curve developed by Zhu *et al.* (2018a).

Reservoir reallocation is simulated by allocating 5% of reservoir storage capacity from the flood control pool to the water conservation pool to enhance water supply performance at the expense of a potential increase in flood risk. This chapter tests reallocation scenarios each month and year-round to illustrate the reallocation strategies under interactions between hydroclimatic and operational factors.

Table 4.1: Investigated Texas reservoirs featured in this chapter. NIDID represents the National Inventory of Dams identification number for the dam.

NIDID	Reservoirs	District	Watershed	Purpose
TX00001	Bardwell	Fort Worth	Trinity	Flood Control, Water Supply
TX00002	Belton	Fort Worth	Brazos	Flood Control, Irrigation, Water Supply
TX00003	Benbrook	Fort Worth	Trinity	Flood Control, Water Supply
TX00004	Canyon	Fort Worth	Guadalupe	Flood Control, Low Flow Augmentation, Water Supply, Fish-Wildlife
TX00005	Grapevine	Fort Worth	Trinity	Flood Control, Water Supply
TX00006	Hords Creek	Fort Worth	Colorado	Flood Control, Water Supply
TX00007	Lavon	Fort Worth	Trinity	Water Supply, Flood Control
TX00009	Navarro Mills	Fort Worth	Trinity	Water Supply, Flood Control
TX00010	Proctor	Fort Worth	Brazos	Flood Control, Water Supply
TX00013	Somerville	Fort Worth	Brazos	Flood Control, Water Supply
TX00014	Stillhouse Hollow	Fort Worth	Brazos	Flood Control, Water Supply
TX00016	Waco	Fort Worth	Brazos	Flood Control, Water Supply
TX00020	Lake O' the Pines	Fort Worth	Cypress	Water Quality, Flood Control, Water Supply
TX00021	Wright Patman	Fort Worth	Sulphur	Flood Control, Water Supply, Low Flow Augmentation
TX04359	Pat Mayse Lake	Tulsa	Red	Flood Control, Water Supply
TX08004	Aquilla	Fort Worth	Brazos	Flood Control, Water Supply, Fish-Wildlife
TX08005	Granger	Fort Worth	Brazos	Water Supply, Flood Control
TX08006	Georgetown	Fort Worth	Brazos	Flood Control, Water Supply
TX08007	Joe Pool	Fort Worth	Trinity	Water Supply, Flood Control
TX08008	Ray Roberts	Fort Worth	Trinity	Flood Control, Water Supply, Water Quality
TX08012	Jim Chapman	Fort Worth	Sulphur	Flood Control, Water Supply

Table 4.2: List of CMIP6 GCMs used in this chapter. Modified from Rastogi *et al.* (2022).

CMIP6 GCM name	Spatial resolution	Ensemble member	GCM institute
ACCESS-CM2	144×192	r1i1p1f1	The commonwealth Scientific and Industrial Research Organization, Australia
BCC-CSM2-MR	160×320	r1i1p1f1	Beijing Climate Center
CNRM-ESM2-1	256×128	r1i1p1f2	French Center National de la Recherche Scientifique
MRI-ESM2-0	160×320	r1i1p1f1	Meteorological Research Institute Japan
MPI-ESM1-2-HR	192×384	r1i1p1f1	The German Climate Computing Center
NorESM2-MM	192×288	r1i1p1f1	Multi-institutional, coordinated climate research project in Norway

4.2.2 Characterizing Hydrologic Extremes and Reservoir Operation Performance

Since most reservoirs in the study area serve both flood control and water conservation purposes, this chapter investigates both flood mitigation and reservoir storage as reservoir operation performance indices. Flood control performance is evaluated by comparing reservoir inflow and release. This chapter defines a flood event as when the streamflow exceeds the 90th percentile of historical data, with a prescribing time lag of at least 10 days between such events. The start and end of an event are marked by the discharge exceeding the threshold before the peak occurrence and falling below the threshold after peak occurrence, respectively. For each flood event, this chapter defines its characteristics including peak discharge (cubic meters per second, cms) and duration (days) (for an illustration see Figure 4.1).

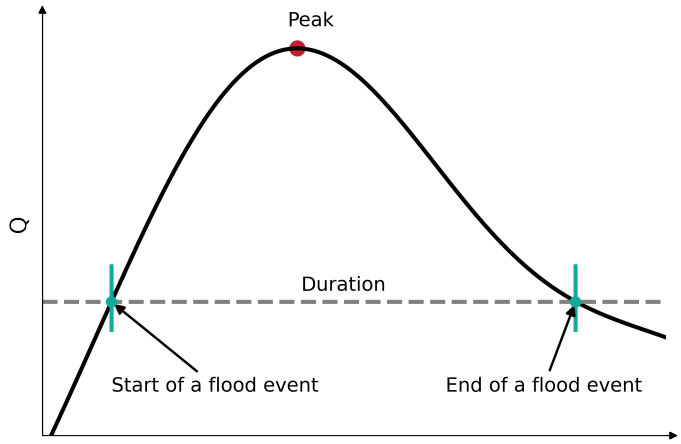

Figure 4.1: Flood characteristics (peak and duration). A flood event is defined as when the inflow exceeds the 90th percentile of historical data, with a prescribing time lag of at least 10 days between such events. The start and end of an event are marked by the discharge exceeding the threshold before the peak occurrence and falling below the threshold after peak occurrence, respectively.

The water supply performance at each reservoir is summarized by the reliability, resiliency, and vulnerability of the reservoir storage time series defined in (Hashimoto *et al.*, 1982). This chapter uses 70% of the conservation pool water supply target, aligning with the Texas Water Development Board's (TWDB) Texas water conditions reports, which label reservoirs as in a "Normal to High" condition when over 70% of their conservation storage capacity is filled. Reliability measures the likelihood that a reservoir can fulfill its water supply target calculated as the percentage of days when reservoir storage is at least 70% of its conservation pool. Resiliency measures how quickly a reservoir recovers from failure states when the storage level is less than 70% of its conservation storage capacity. Vulnerability quantifies how severe the consequences of failure may be (how unfilled the reservoir is), calculated as the proportion of the shortfall (amount of water required to reach 70% of the conservation pool) to the total capacity of the conservation pool.

The trained data-driven models (representing the existing operation policy) are driven by each projected climate forcings in the historical period (1980-2014) to represent baseline

performance and future period (2015-2059) to represent future performance. This allows the evaluation of reservoir performance under consistent climate change scenarios.

4.2.3 Structure Equation Modeling (SEM) for Reservoir Storage Shortages under High Inflow Conditions

This chapter utilizes Structural Equation Modeling (SEM) to disentangle the complex relationships among hydrologic variability, operation strategies, evaporation, and reservoir storage shortages. As a robust multivariate statistical tool, SEM enables the simultaneous examination of multiple observed and latent variables (Kline, 2023). It combines factor and path analysis for comprehensive cause-effect modeling (Bollen and Pearl, 2013). In our SEM model, the author establishes 'StorageDrought' as a latent variable, measured by three observed metrics: the frequency, duration, and magnitude of occasions when reservoir water levels fall below 70% of the conservation capacity. This chapter particularly focuses on the role of high flow (defined as flows exceeding the 70th percentile of historical inflow) in potentially mitigating storage drought. For this, the author introduces 'HydroCondition' as another latent variable, measured by the frequency, duration, and magnitude of high inflow events. Reservoir management practices during flood seasons aim to conserve water and lessen storage drought risk. The author thus introduces a third latent variable, 'OperationStrategy', measured by the frequency, duration, and magnitude of water-release events. Reservoir evaporation is included as a factor exacerbating storage drought, with all variables analyzed on a monthly scale.

4.3 Results

The trained reservoir operation models demonstrate robust performance across most reservoirs investigated in this chapter, as detailed in Figure 4.2. The median NSE values on the test set for reservoir storage (or reservoir water level), evaporative loss, and water

release are 0.991, 0.984, and 0.938, respectively. To evaluate the performance of the current reservoir operation policy in terms of flood control and water supply under future climate, this chapter runs the trained reservoir model with baseline and future streamflow from the downscaled hydroclimatic models.

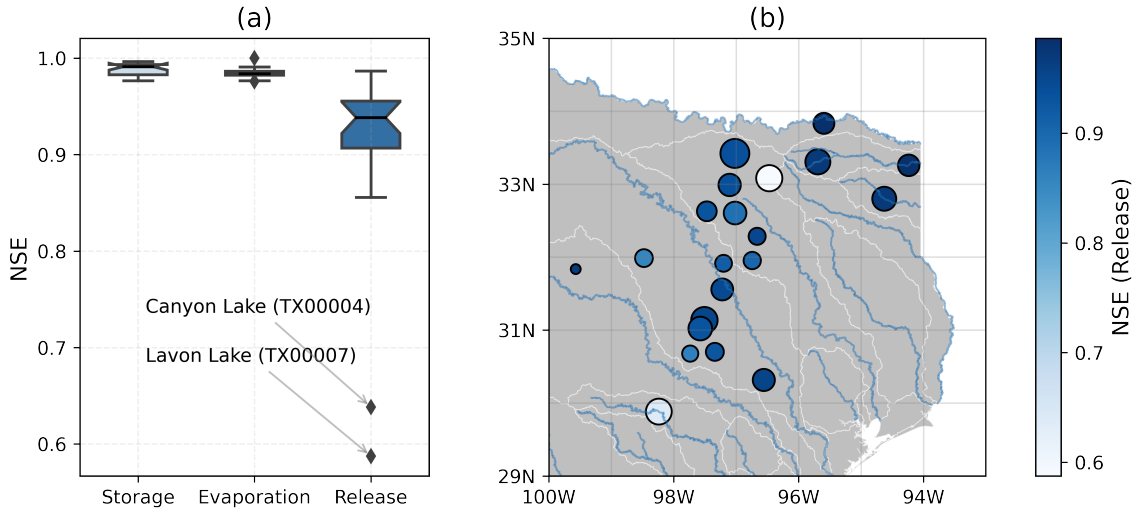

Figure 4.2: (a) Performance of reservoir operation models. The median NSE values on the test set for reservoir storage (or reservoir water level), evaporative loss, and water release are 0.991, 0.984, and 0.938, respectively. (b) Spatial distribution of NSE for the selected reservoirs in this chapter.

4.3.1 Performance of Current Reservoir Operation Policies for Flood Risk Mitigation and Water Supply under Future Climate

The change of high flow conditions under future climate is evaluated based on the differences between future (2020-2039 and 2040-2059) and baseline (1990-2019) periods' reservoir inflows. There is a modest change (within a range of +/- 20%) for most reservoirs in terms of the number and duration of flood events (Figrue 4.3a) compared to the baseline period. However, all reservoirs are projected to experience larger flood peaks (rising by between 2% and 62%), indicating a higher risk of more severe flood events upstream in the future.

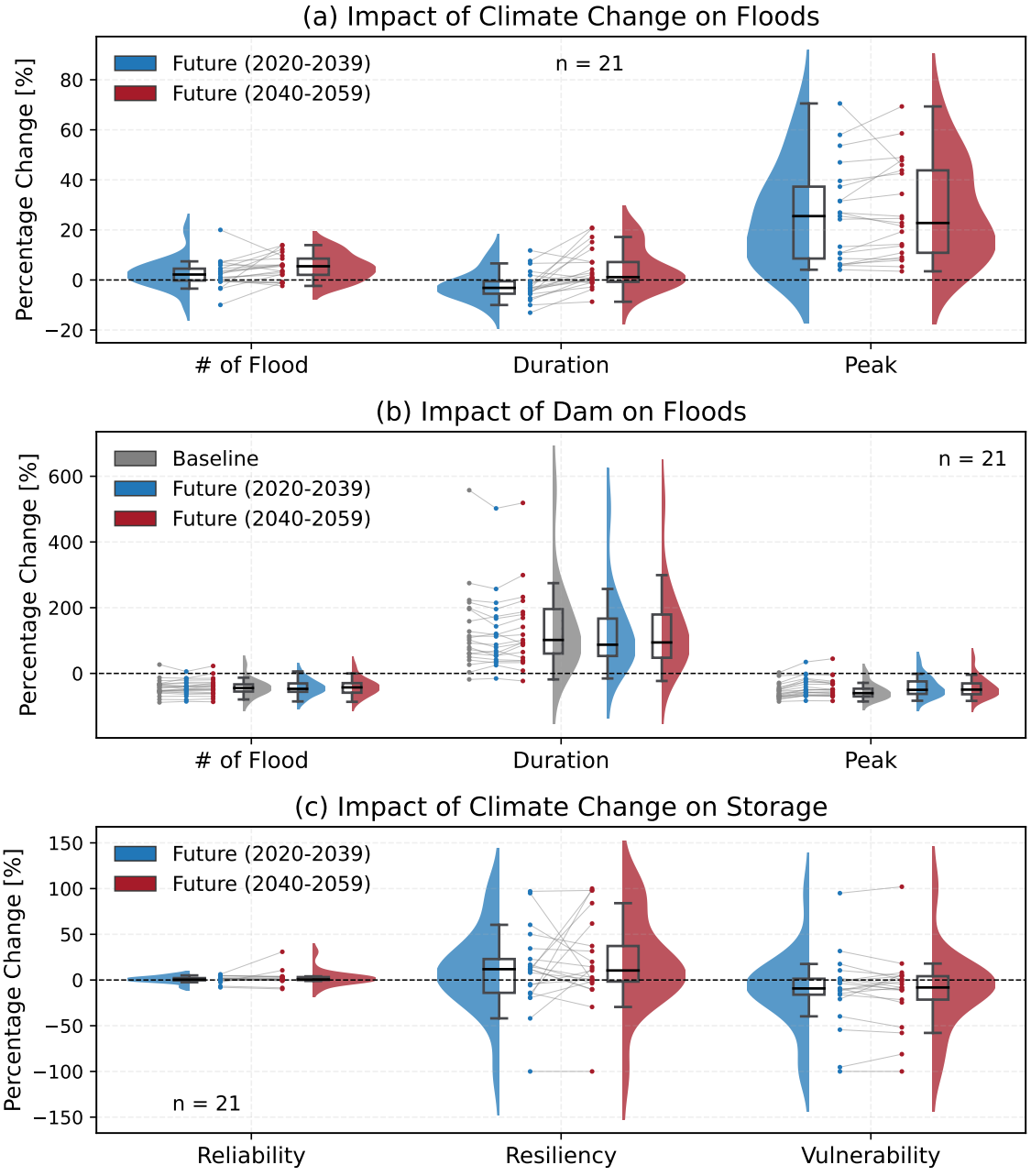

Figure 4.3: (a) Changes in flooding indices calculated from reservoir future inflows relative to baseline period; (b) Percentage changes in indices in flooding indices from reservoir inflow and release during baseline and future period; (c) Percentage change in storage indices attributable to climate change. The semi-violin plots illustrate the distribution of percentage changes, while the semi-box plots indicate median values, as well as the first and third quartiles (depicted by upper and lower box hinges). Whiskers show the range of maximum and minimum values, excluding outliers. Individual dot points represent raw data for each reservoir. "n=21" represents the number of selected reservoirs.

This chapter compares the reservoir inflow and release to quantify the changes in flood risks to evaluate the reservoir's performance in mitigating flood risks. Existing operation rules are largely effective in reducing flood risks across the baseline, near-future, and mid-term future periods (Figrue 4.3b). Reservoirs manage to decrease the frequency of flood events (around -10% to -80% across all 21 reservoirs for all periods) and reduce the magnitude of flood peak flows (around -2% to -90%) by eliminating smaller floods and reducing the magnitude of flood peak flows. The flood water is temporally saved in the reservoir and released after the event, as indicated by extending the duration of events. However, Pat Mayse Lake stands as an exception. It may be attributed to the largest shifts in flow timing in the future (as detailed in Supplementary Materials Figure 4.4), specifically a higher concentration of water flow in the flood season, particularly in the summer. This could also highlight potential limitations in current management strategies for future conditions.

The impact of climate change on the water supply performance differs across reservoirs (Figrue 4.3c). In the near future (2020-2039), water supply reliability will experience mild changes (within ±15%), a trend poised to persist and amplify in the mid-term (2040-2059), potentially benefiting from increasing incoming water quantities. The reservoir in Texas tends to maintain the water level at the designated top of conservation pool elevation as stream flows and water demands allow (Wurbs, 2021). Minor changes in reliability are expected and align with the purposes of these reservoirs to provide a consistent water supply. While reliability remains fairly stable, the resiliency and vulnerability of the 21 reservoirs are noticeably affected. In the near and mid-term future, under unchanged operational rules, some reservoirs are projected to experience increased resiliency and decreased vulnerability. However, not all reservoirs are resistant to future droughts. Some, like Canyon Lake and Bardwell Lake, face a bleak future with declining resiliency and heightened vulnerability, leading to severe failures and recovery difficulties. It would be advantageous for managers to revisit and update the existing operation rules to better equip

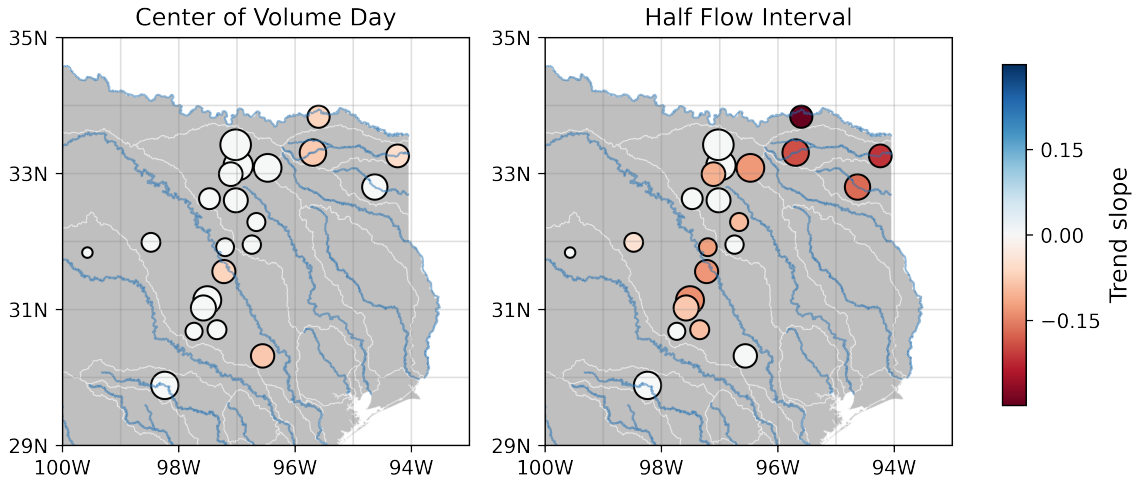

Figure 4.4: Trends in the spatial distribution of two metrics for the selected 21 reservoirs: the center of volume day and the half flow interval. The center of volume day refers to the day of the year by which half of the annual streamflow has passed. A decreasing trend in this metric suggests that the midpoint of the annual streamflow is occurring earlier in the year, meaning that a larger volume of water is flowing during the first half of the year. The half flow interval is defined as the time span between the day a quarter of the year's total streamflow has passed, and the day three quarters of the annual streamflow has passed. A decreasing trend in this interval indicates that this 50% of the total streamflow is happening over a shorter time period, pointing to a higher concentration of water flow in the flood season, particularly in the summer.

themselves against the anticipated higher risk of extreme events and potential degradation in water supply performance.

4.3.2 Impact of Feedback between Operation Strategy and Hydrologic Condition on Storage Drought and Evaporative Loss

The reservoir storage droughts shown in Figure 4.3c would be caused by intertwined factors including inflow regime, evaporative loss, and reservoir release. For example, more frequent high flows would lead to spilled flood water, which decreases the water stored in the reservoir even with more inflow. High water levels under the conservation operation target lead to a larger reservoir surface area and enhanced evaporative loss, which may result in reservoir storage drought. The SEM analysis provides quantitative insights from the

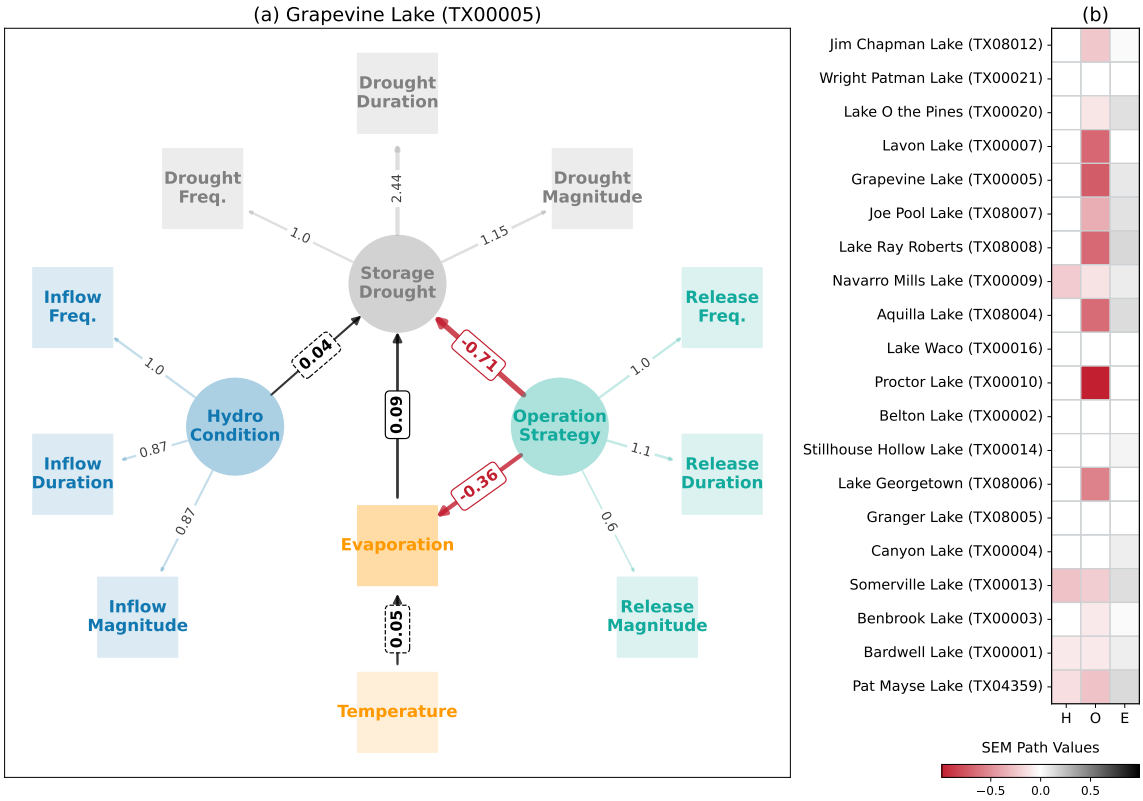

Figure 4.5: (a) Influence of hydrologic conditions, operation strategies, and evaporation on storage drought at Grapevine Lake, as evaluated by a Structural Equation Model (SEM). Latent variables are represented by ellipses, while observed variables appear in rectangles. Standardized path values are displayed along each pathway; positive values are enclosed in black boxes and negative values are in red boxes. Paths with a statistically significant p-value below 0.05 are enclosed in a solid line box, while those with nonsignificant p-values are enclosed in a dashed line box. (b) Summary of SEM path values for select reservoirs investigated in this chapter. In the x-axis labels, "H" represents hydrologic conditions, "O" denotes operation strategies, and "E" signifies evaporation. Statistically insignificant values are shown as Not a Number.

machine learning reservoir model to attribute the hydroclimatic and operational factors on reservoir storage drought. The SEM analysis implies that current management practices adopted in Texas generally help in mitigating the risks of water shortages, thus facilitating a stable water supply, while evaporation emerges as a deteriorating factor for the system performance in many reservoirs. For example, in the case study of Grapevine Lake (Figure 4.5a), operation strategies for managing high inflow effectively alleviate storage drought

(with a path value of -0.71); conversely, evaporation exacerbates storage drought (with a path value of 0.09). Herein the path value measures the direct impact of one variable on another within the model, where its magnitude reflects the strength of this effect and its sign (positive or negative) denotes the nature of the relationship, similar to a regression coefficient in linear regression. Figure 4.5b further summarizes the quantitative influence of hydrological conditions, operation strategies, and evaporation on storage drought of selected reservoirs. For most reservoirs, the current operation strategies stand as the paramount factor in reducing storage drought occurrences. However, there are exceptions; a few reservoirs either do not exhibit a statistically significant relationship with any of the factors examined (e.g., Lake Waco) or show a higher correlation with other factors like evaporation (e.g., Canyon Lake) or hydrological conditions (e.g., Navarro Mills Lake). This may indicate that operation strategies may be less effective in these reservoirs, requiring an update of the current regulation rules to better capitalize on high inflow conditions.

The non-beneficial reservoir evaporative loss represents the water cost of reservoir operation and will affect the tradeoffs among various operation targets. In a specific reservoir, meteorological conditions and the reservoir water level (or surface area) jointly influence the evaporative process over seasons. Monthly correlation analysis identifies dominant factors influencing reservoir evaporation across seasons under the six climate-forcing scenarios, as shown in Figure 4.6. The author categorizes the reservoirs into three groups based on the dominant influencing factors: 1) storage-dominated (such as the Pat Mayse Lake), where the water storage level is the principal determinant of evaporation; 2) temperature-dominated (such as the Jim Chapman Lake), where temperature mainly governs evaporation all year round; 3) seasonally-hybrid (such as the Lake Waco), where storage and temperature alternately have a significant influence over different seasons. The different dominant factors on reservoir evaporative loss also provide insights for adaptive reservoir operation in a changing climate. For seasonally-hybrid reservoirs, temperature is the primary control from

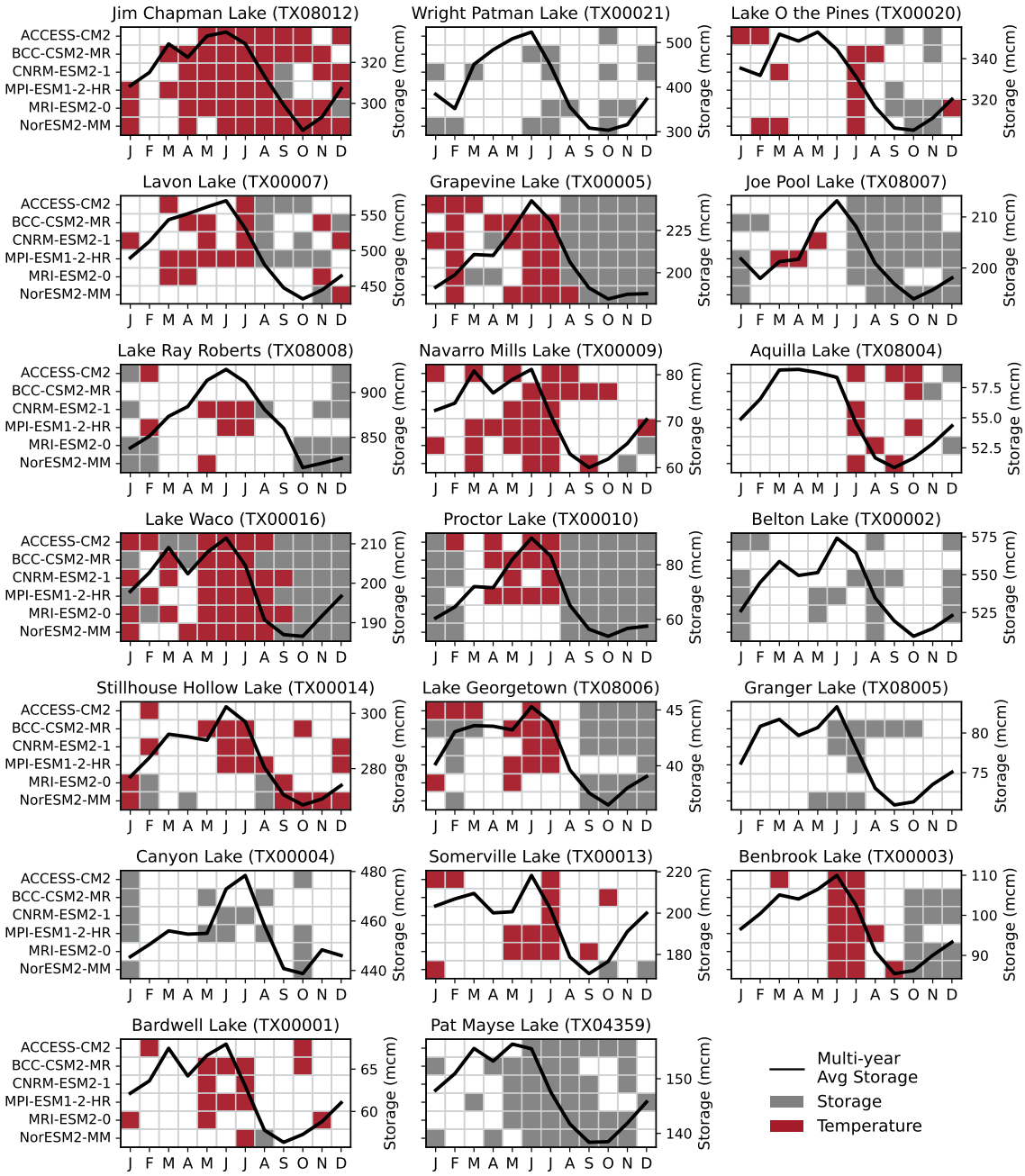

Figure 4.6: Influential determinants of reservoir evaporation, segmented by month and informed by six GCMs, for the selected reservoirs examined in this chapter. Gray indicates that storage is the dominant factor affecting evaporation for the corresponding month, while red means that temperature plays a more significant role. The black solid line displayed on the secondary axis represents the multi-year average storage.

May to August, which implies that in summer, curbing evaporation by managing water levels is largely unfeasible. Conversely, from September to December, usually the post-flooding season, water level becomes a more important factor influencing evaporation. Therefore, any decision to alter the conservation level for storing more water during these months should be approached with caution to avoid inadvertently increasing the evaporative loss.

4.3.3 Opportunity and Strategies of Reservoir Reallocation to Tradeoff between Water Supply and Flooding Control

Based on the analysis in Sections 4.3.1, and 4.3.2, reallocation strategies by adjusting the existing conservation pool's upper limits may prove additional benefits for water supply performance without substantially raising the risk of downstream flooding and evaporation loss. For example, for Lake Waco (Figure 4.7a), elevating the year-round conservation capacity enhances the reservoir's resiliency, with slight improvements in reliability and vulnerability. This positive effect is notably observed in August, around the end of flood seasons. Increasing the conservation pool level for the whole year only yields a slightly better performance than the August adjustment. The adjustment in August enables the harnessing of more floodwater during the final peak flows, subsequently delaying the onset of storage recession below the original top of the conservation pool (Figure 4.7c). It facilitates a more effective recovery from storage drought (i.e., the water level falling below 70% of the conservation pool). Meanwhile, the elevated conservation pool does not significantly escalate the threat of downstream flooding, with fluctuations remaining within a +/- 0.2% range, highlighting the feasibility of adjusting the conservation capacity on a monthly or annual basis (Figure 4.7b). Figures 4.9 and 4.10 in Supplementary Materials demonstrate the feasibility of reallocation in other reservoirs, with nuanced differences among cases. The optimal strategy for Lake Waco and Pat Mayse Lake is to temporarily elevate the designated top of the conservation pool in August; however, Jim Chapman Lake benefits

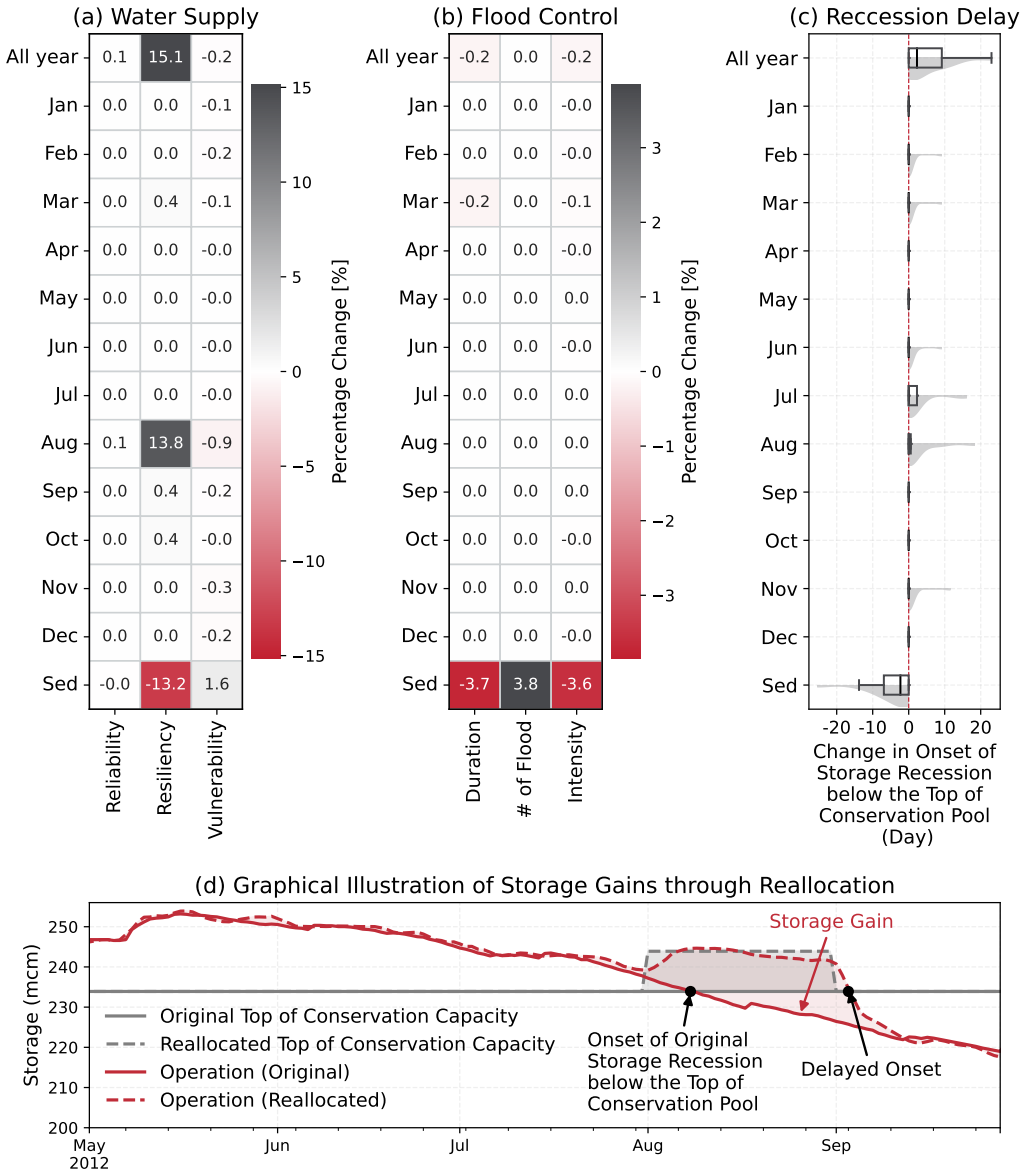

Figure 4.7: (a) Percentage change in water supply performance (reliability, resiliency, vulnerability) in different storage reallocation scenarios in Lake Waco. (b) Percentage change in downstream flood risk (duration, frequency, intensity) in different storage reallocation scenarios in Lake Waco. (c) Changes in the onset of storage recession in different storage reallocation scenarios in Lake Waco. The onset of storage recession is defined as the first day the reservoir's storage falls below the conservation level after a prolonged period of sustaining at this level, and this recession should last for at least 14 days. (d) Hydrograph of reservoir storage fluctuations between May 1st and September 30th, 2012. "All Year" refers to a scenario involving year-round elevation of the conservation pool's upper limit. Labels from "Jan" to "Dec" indicate scenarios where the pool elevation is raised only during the corresponding month. "Sed" represents a condition where sedimentation decreases available storage.

from a February adjustment, given its additional flood events during the winter months. This adjustment for Jim Chapman Lake allows the reservoir to retain more floodwater following winter flood events, with slightly reducing the flooding risk downstream (Supplementary Materials Figure 4.10).

Quantifying the tradeoff between high inflow harvesting and evaporative losses is vital for reallocation strategies to safeguard reservoir efficiency and stable water supply. The water saving from reallocation would be counterbalanced by the high evaporation loss, leading to negligible increases in the water supply performance. Taking Lake Waco as an exemplar, compared to Figure 4.7a, supplementary Materials Figure 4.8 reveals that the elevated conservation capacity level in September is counteracted by an increase in reservoir evaporation. For Lake Waco (seasonally-hybrid type), September is a critical period following the flood season, where the water storage predominantly controls the reservoir evaporation, overshadowing the influence of temperature (see Figure 4.6). Therefore, the analysis of dominant factors in Figure 2 and their temporal pattern in Figure 4.6 provides a quantitative simulation approach to develop reservoir reallocation strategies by considering the tradeoffs and synergies among the temporal characteristics of hydroclimatic and operational factors. It is noted that sedimentation (such as reduction in active reservoir storage) stands as an additional threat to reservoir operation performance besides evaporation. For Lake Waco, the reservoir sedimentation-induced capacity loss tends to adversely affect both resiliency (-13.2%) and vulnerability (+1.6%). This drawback can negate the benefits offered by reallocation strategies, hinting at the potential merits of sediment removal for the system's sustainability (Kawashima, 2007). While sedimentation slightly elevates the flooding risk downstream by intensifying its frequency, this increment is relatively minor, recorded at +3.8%.

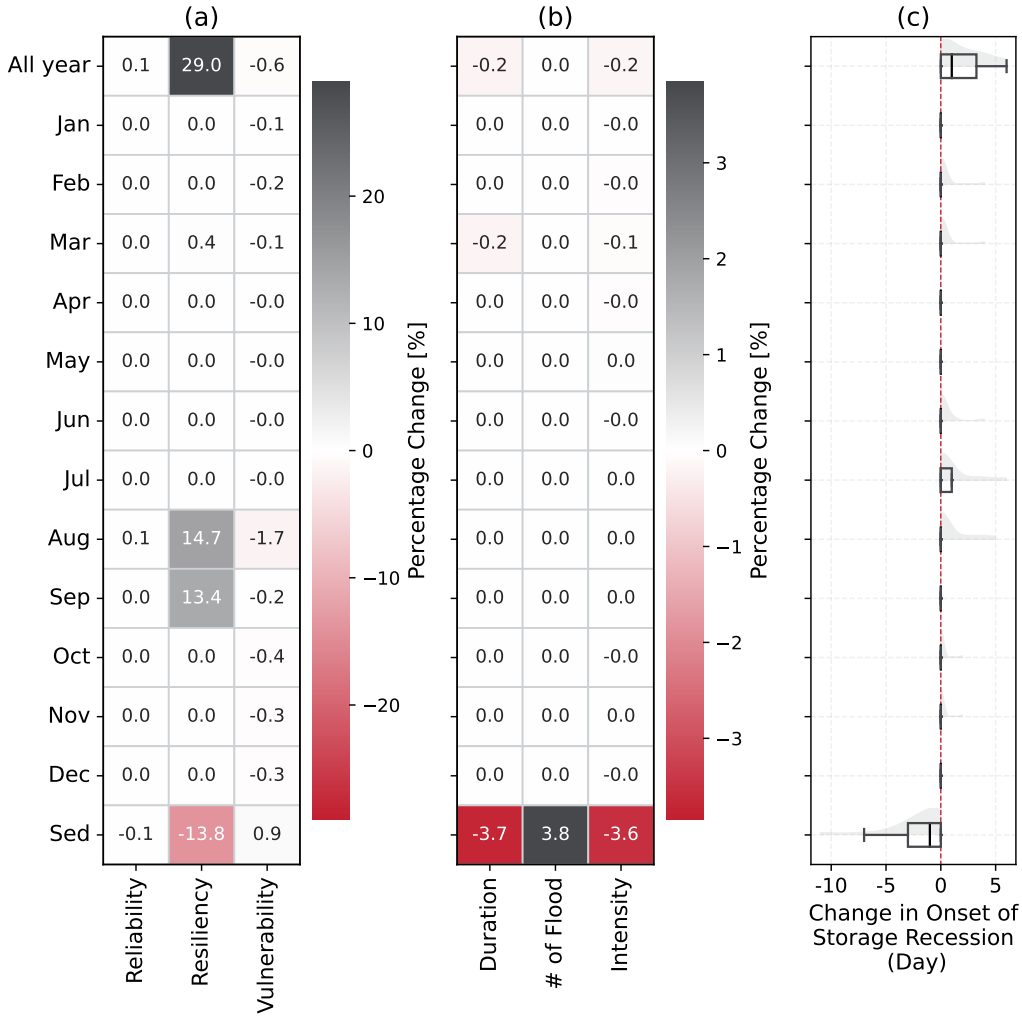

Figure 4.8: (a) Percentage change in water supply performance (reliability, resiliency, vulnerability) in different storage reallocation scenarios (without considering the impact of evaporation) in Lake Waco. (b) Percentage change in downstream flood risk (duration, frequency, intensity) in different storage reallocation scenarios in Lake Waco. (c) Changes in the onset of storage recession in different storage reallocation scenarios in Lake Waco. The onset of storage recession is defined as the first day the reservoir's storage falls below the conservation level after a prolonged period of sustaining at this level, and this recession should last for at least 14 days. "All Year" refers to a scenario involving year-round elevation of the conservation pool's upper limit. Labels from "Jan" to "Dec" indicate scenarios where the pool elevation is raised only during the corresponding month. "Sed" represents a condition where sedimentation decreases available storage.

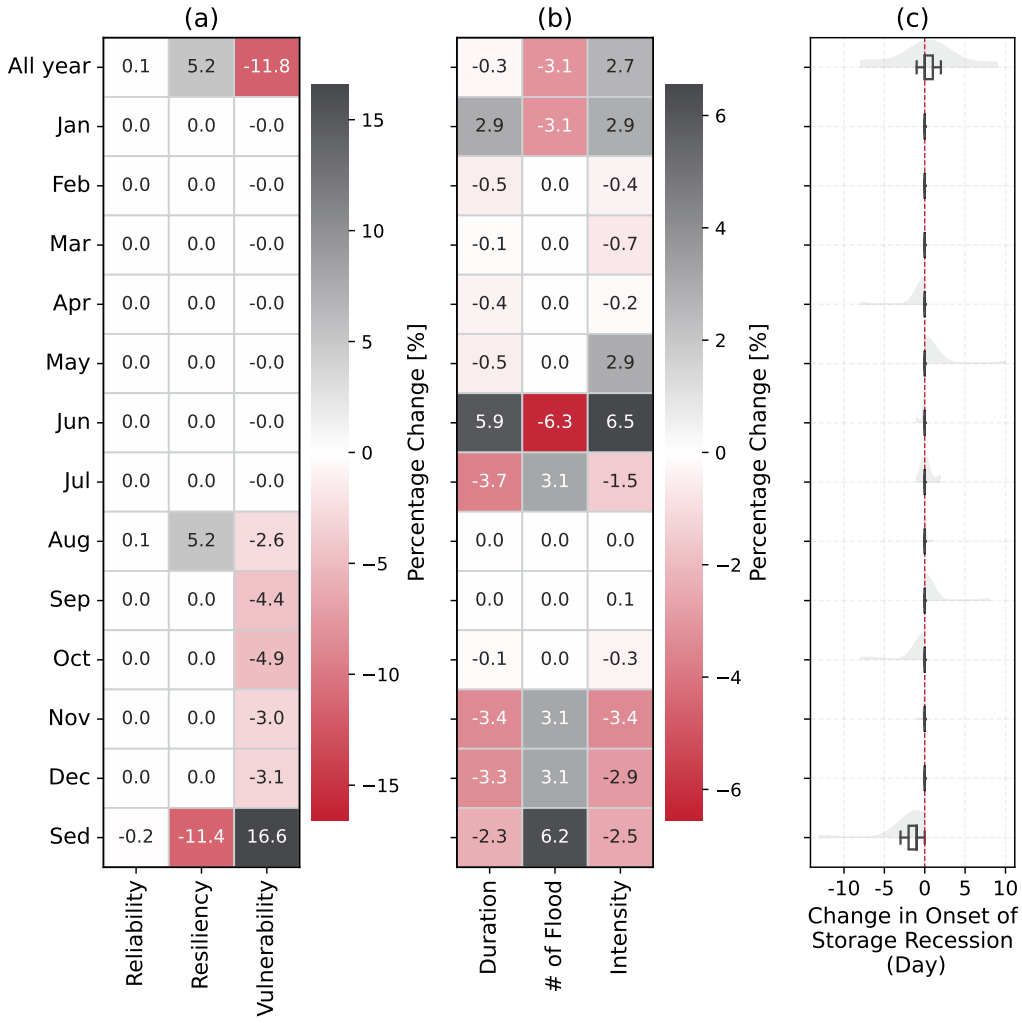

Figure 4.9: (a) Percentage change in water supply performance (reliability, resiliency, vulnerability) in different storage reallocation scenarios (considering the impact of evaporation) in Pat Mayse Lake. (b) Percentage change in downstream flood risk (duration, frequency, intensity) in different storage reallocation scenarios in Pat Mayse Lake. (c) Changes in the onset of storage recession in different storage reallocation scenarios in Pat Mayse Lake. The onset of storage recession is defined as the first day the reservoir's storage falls below the conservation level after a prolonged period of sustaining at this level, and this recession should last for at least 14 days. "All Year" refers to a scenario involving year-round elevation of the conservation pool's upper limit. Labels from "Jan" to "Dec" indicate scenarios where the pool elevation is raised only during the corresponding month. "Sed" represents a condition where sedimentation decreases available storage.

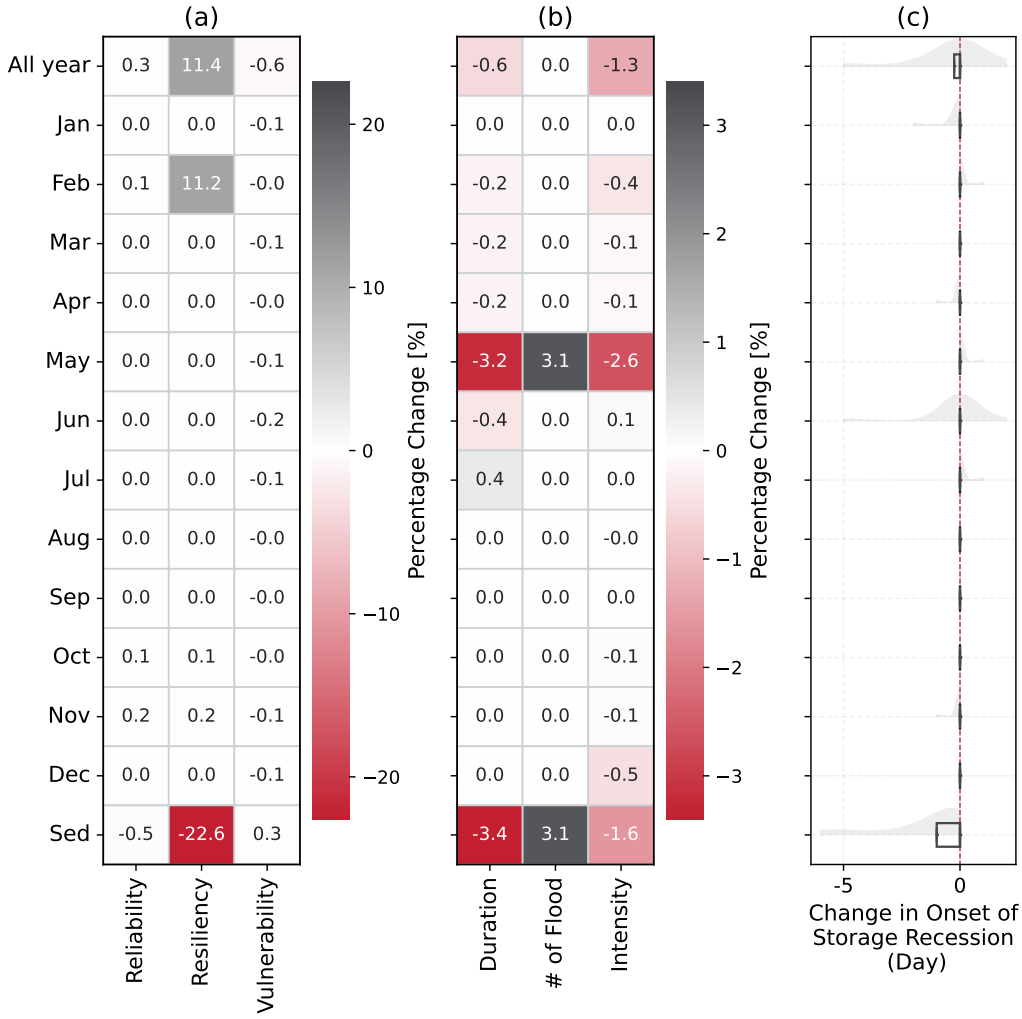

Figure 4.10: (a) Percentage change in water supply performance (reliability, resiliency, vulnerability) in different storage reallocation scenarios (considering the impact of evaporation) in Jim Chapman Lake. (b) Percentage change in downstream flood risk (duration, frequency, intensity) in different storage reallocation scenarios in Jim Chapman Lake. (c) Changes in the onset of storage recession in different storage reallocation scenarios in Jim Chapman Lake. The onset of storage recession is defined as the first day the reservoir's storage falls below the conservation level after a prolonged period of sustaining at this level, and this recession should last for at least 14 days. "All Year" refers to a scenario involving year-round elevation of the conservation pool's upper limit. Labels from "Jan" to "Dec" indicate scenarios where the pool elevation is raised only during the corresponding month. "Sed" represents a condition where sedimentation decreases available storage.

4.4 Discussion

4.4.1 Climate Change Uncertainty

The uncertainty in decision-making regarding future hydroclimatological outcomes arises from various sources: SSP scenarios, GCMs, downscaling methods for local weather conditions from coarse-resolution simulations, bias-correction techniques, and hydrologic models (Nielsen-Gammon *et al.*, 2020). Focusing on the SSP585 scenario, which is the most severe climate projection with an anticipated additional radiative forcing of 8.5 W/m^2 by 2100, this chapter examines its implications for reservoir management in mitigating flood risks and securing water supply under extreme climate conditions. Characterized by a high reliance on fossil fuels and resource-heavy lifestyles (Masson-Delmotte *et al.*, 2021), SSP585 presents pressing challenges for water resource management. Expanding the scope, other scenarios such as SSP126 ("Sustainability", adding 2.6 W/m^2 radiative forcing by 2100), SSP245 ("Middle of the road", with 4.5 W/m^2), SSP370 ("Regional rivalry", reaching 7 W/m^2), and SSP460 ("Inequality", with 6 W/m^2), cover a spectrum of possible future conditions for water management. Further investigation under these scenarios can provide a broader probability range with different levels of climate protection measures, spanning from optimistic to more moderate future outcomes. However, the efficacy of such investigations hinges on the reliability of climate models. The six GCMs selected for this chapter have relative independence and low structural similarity, which should enhance the degree of confidence in climate projections (Pathak *et al.*, 2023). Despite this, as noted by Ashfaq *et al.* (2022), these models have limitations in simulating seasonal mean precipitation and temperature for the South and West U.S. regions. Furthermore, although dynamic and statistical downscaling methods are similarly effective in correcting historical climate simulations, the projected precipitation extremes based on Livneh (Livneh *et al.*, 2015) as the reference data tend to be underestimated (Rastogi *et al.*, 2022), potentially

leading to underestimated flood risk assessments in our study. Additionally, bias correction, which primarily adjusts the monthly distribution of precipitation and temperature, may overlook interannual variability (Kao *et al.*, 2016). Coupled with the potential biases in hydrologic modeling due to variations in spatial resolution and parameterization (Kao *et al.*, 2022), the uncertainty in hydrological outcomes may be further compounded. This highlights the importance of cautious interpretation in applying these models to water resource management, considering the uncertainties inherent in climate change projections.

4.4.2 Reservoir Operation Targets

Evolving operational targets and water demand can present a mix of opportunities and challenges for future reservoir management. This chapter assesses whether current reservoir operation rules can continue to meet their targets amid a changing climate. It is assumed that the designated operational purposes of the reservoirs will remain unchanged in the future. However, these purposes may evolve, either realigning operational priorities or expanding to include new objectives (e.g., adding a hydropower component for hydroelectric generation, or allocating environmental flow for downstream ecosystem protection). Hadjerioua *et al.* (2012) highlight that several non-power dams in Texas (including Belton Lake, Livingston Dam, etc.) have the potential to generate significant amounts of clean, reliable hydropower, with the estimated annual average generation ranging from 59,153 to 327,233 MWh. As a case in point, Livingston Dam underwent modifications by adding three 8-MW Kaplan-type turbine-generator units, which began commercial operations on July 14, 2020, and now produce sufficient clean electricity to supply around 12,000 households in East Texas (Chase-Israel, 2022). Such modifications could be implemented in other dams in the future as well, potentially leading to a shift in the original purposes and operational strategies of these reservoirs. While this could ease pressure on the power system, it might also heighten the trade-offs between hydropower generation and flood control (Opperman *et al.*, 2022).

Due to data and model limitations, this chapter refrains from directly modeling specific demands such as crop requirements or municipal and industrial (M&I) water demand. Instead, the author assumes that these aspects are implicitly embedded in the historical records, which have guided operators to meet water supply targets in the past. In fact, per capita water usage does not always align predictably with population growth, making it challenging for data-driven models to reasonably represent these complex dynamics. An example of this is seen in Texas, where remarkable water conservation efforts have led to a notable decrease in average municipal water consumption per person, from 175 gallons daily in 2000 to 138 gallons in 2015 (a reduction of around 21%), as reported by the Texas Water Development Board. Continuing water conservation efforts in the future could help alleviate stress on water supply systems.

A caveat is that this chapter does not account for the evolving nature of water rights, regulations, and laws, which can change in response to various environmental and societal factors. It is promising for future work to incorporate details from region-specific water management models, such as the Water Availability Model (WAM, Wurbs *et al.*, 2005), which contains the Water Rights Analysis Package (WRAP) executable programs for individual Texas river basins.

4.4.3 Forecast-informed Reservoir Operation (FIRO)

Effective management strategies to enhance the water supply performance may further benefit from incorporating accurate forecasting information. The storage reallocation strategies examined in this chapter are based on historical data, overlooking real-time meteorological and hydrological forecasts. In the future, the actual flood season patterns may deviate from historical expectations. This deviation could pose water management challenges, such as unanticipated flood damages (i.e., if the flood season is delayed and the actual flood season is longer than a regular one, a flood event occurs beyond a regular

flood season, Ding *et al.*, 2015) or prolonged droughts (i.e., if the actual flood season ends earlier than a regular one, the dry season will be extended). Forecast-informed reservoir operation (FIRO) is a management strategy that leverages watershed monitoring data along with advanced weather and streamflow forecasts to aid water managers in making informed decisions based on current and anticipated conditions (Jasperse *et al.*, 2020; Delaney *et al.*, 2020). Short-term forecasts are informative for a water reservoir system primarily operated to fulfill the short-term objective (e.g., flood control). They equip the system operators with the foresight needed to prepare buffer storage for mitigating the upcoming flood peak and thus minimizing flood damages (Wang *et al.*, 2012; Raso *et al.*, 2014; Zhao *et al.*, 2014). Medium to long-term forecasts offer valuable insights for reservoirs serving the purpose of water supply, which is typically characterized by slower dynamics that unfold over extended periods (Denaro *et al.*, 2017). Recent initiatives in FIRO have demonstrated the potential of utilizing forecasted rainfall and real-time reservoir inflow data to optimize water management (Konieczki *et al.*, 2017; Delaney *et al.*, 2020; Zarei *et al.*, 2021). By taking advantage of accurate forecasts, the reservoir system in Texas can more effectively utilize the last flood event of each season to replenish the conservation pool instead of releasing water downstream, thereby enhancing water storage benefits. This strategy can simultaneously meet flood control objectives and water supply needs.

While FIRO is viewed as a cost-effective method for ensuring water supply and reducing flood impacts, its adoption in Texas is limited due to specific challenges, including limitations in forecast accuracy and specificity, infrastructural and operational constraints facing reservoir operators, and information gaps between forecast providers and operators (Fernando *et al.*, 2020). The expansive geographical scale and climatic diversity of Texas present a huge challenge for accurately predicting storm systems that lead to runoff and flooding over extended timeframes (Texas Water Development Board *et al.*, 2020). In Texas, the unpredictability of precipitation is substantially influenced by weather patterns like El

Niño and La Niña (Hoerling *et al.*, 2013; Cheng *et al.*, 2018), and further compounded by the erratic nature of thunderstorms and tropical disturbances during the warm season (Nielsen-Gammon *et al.*, 2020). This unpredictability particularly impacts streamflow forecasts, where the accuracy in predicting high flows sensitively depends on the quality of the precipitation forcing (Kim *et al.*, 2018). Perhaps out of a risk-averse inclination, many reservoir operators in Texas may typically rely on absolute statements, not probability. Yet, constraints from infrastructure and operations demand skillful streamflow predictions. For those non-federal water supply reservoirs in Texas, which secure municipal and industrial water use in large cities, they lack a designated flood control pool and have limited capacity to regulate, store, and manage floodwaters. Therefore, pre-releases from conservation pools before potential flood events rely heavily on the accuracy of forecasts. To promote the adoption of FIRO, it is critical on one hand to ensure reliable forecasts for effective reservoir management, maintaining steady revenue and a continuous water supply for consumers. On the other hand, it is imperative to determine how to modify operation strategies to counteract the risks posed by climate change in these reservoirs. Furthermore, infrastructural challenges, such as the design of reservoirs for water conservation and the potential need for retrofitting or adding additional flood gates, must be tackled to fully harness the benefits of proactive forecast usage (Fernando *et al.*, 2020).

4.5 Conclusions

This chapter evaluates the efficacy of current operational rules in addressing future flood risks and water supply performance. While current practices effectively counter many flood risks, climate change amplifies challenges in reservoir resilience and vulnerability. In Texas, most operational strategies adeptly manage high inflow conditions, mitigating storage drought and ensuring a reliable water supply. Yet, certain reservoirs may not benefit sufficiently from these guidelines in the future, emphasizing the need for regulatory updates.

The adverse impact of reservoir evaporation on the performance of many water supplies should be factored into these revisions.

This chapter further analyzes the potential benefits of adjusting the upper limits of reservoir conservation pools. Our findings highlight the feasibility of storage reallocation between the flood control and conservation pools to bolster water supply resilience without amplifying downstream flood risks. This straightforward way can pave the way for broader FIRO adoption, which in turn offers more accurate dynamic storage adjustments to store more floodwater. Moreover, reservoir sedimentation reduces resilience and heightens vulnerability, underscoring the importance of sediment removal in ensuring sustainable operations.

Chapter 5

CONCLUSIONS AND PERSPECTIVES

5.1 Conclusions and Implications

This dissertation demonstrates the enhancement of hydrologic models through advanced techniques and evolving datasets, aimed at improving their accuracy and utility. The first part of work (Chapter 2) tackles the prevalent issue of missing or inadequately represented components in the existing large-scale hydrologic models. Taking the simulation of regulated flow conditions as an example, the study presents a framework for all data-driven models to utilize limited data sources more effectively. Specifically, this dissertation employs ML and multiple temporal scale information from the observed reservoir operation datasets, thereby improving the representation of reservoir operations for at least 60% cases, especially for the situation where the key explanatory variables (i.e., inflow or water level) are not available. Moreover, Chapter 2 explores the interpretability of ML models by utilizing sensitivity analysis (e.g., the permutation feature importance in this study) to determine which types of hydrologic information (e.g., inflow, water level, snow depth) predominantly influence the decision-making process for flow regulations across various temporal scales (e.g., daily, monthly). The second part (Chapter 3) addresses the computational challenges associated with model parameterization. The development of a ML-based surrogate model expedites the parameter estimation and calibration. This computational demand arises not just during initial model calibration but also in scenarios where remotely sensed parameter information is integrated into the model (e.g., determining which remote sensing products is most appropriate) and during re-calibration after updating any parameter information. A surrogate model of the process-based model can also offer hydrologists and stakeholders a flexible

platform for rapid scenario exploration and hypothesis testing. Additionally, the scientific insights offered by this chapter include, increasing physical realism of hydrologic model calls for the adoption of dynamic parameterization, particularly for those properties that may vary over time. Vegetation parameterization serves as a case study in this dissertation, illustrating how the dynamic nature of vegetation can influence hydrologic responses. It further suggests that the assessment of climate change on the water cycle needs to consider the feedback between climatic conditions and vegetation dynamics as vegetation can evolve as the climate changes. From the perspective of hydrologic modelers, these two parts of work (Chapter 2 and 3) enhance the hydrologic model's realism by improving both its representation and parameterization, respectively. For the water managers, a combination of surrogate model (that can be used for naturalized runoff simulation) and reservoir operation module (that receives the upstream flow conditions as input) enables the simulations of reservoir behaviors under different climate projection scenarios. The last part of work (Chapter 4) shows such application to translate hydrologic and climatic data into actionable strategies for water management. Using 21 federal reservoirs in Texas as a testing case, this chapter provides a framework for stakeholders to evaluate the effectiveness of the current reservoir operation policy under future climate through the interactions among hydroclimatology, reservoir infrastructure, and operation policy.

The primary contribution of this dissertation lies in demonstrating how the integration of sophisticated quantitative methods and diverse data sources can aid in addressing scientific challenges in hydrologic modeling and water management. It is strategically aimed at guiding hydrologists and water managers in the application of these tools. The novelty of this work is encapsulated in three general scientific findings: (1) for data-driven models, the prediction skills and forecasting ability could benefit from the comprehensive utilization of data information, especially for data-scarce situations (Chapter 2); (2) in process-based models, certain parameters that may evolve alongside other parameters or processes require

dynamic parameterization (e.g., the incorporation of the real-time remotely sensed dataset and the recalibration upon updating parameter information) for accurate assessment in the context of climate change (Chapter 3); (3) in water management, operational costs should be carefully weighed, as it is vital to balance costs and benefits when implementing changes in water management strategies (Chapter 4).

5.2 Future Work

5.2.1 Explicitly Encoding the Physical Process in the ML models to Facilitate Scientific Understanding

From a modeler's perspective, a key distinction between ML models and process-based models is found in the balance between parameters and structure—a representation of the underlying processes. ML models (inherently data-driven) typically feature relatively simple formulas (e.g., the most basic expression in neural networks is a linear transformation to the input vector through a weights matrix, the non-linear activation functions may or may not be applied to the results of linear transformations) accompanied by a large number of parameters (e.g., weights and biases in neural networks). This often leads to criticisms of these models being over-parameterized (Rocks and Mehta, 2022). In contrast, process-based models are often characterized by well-defined structural expressions (while not universally accepted as entirely accurate, Clark *et al.*, 2017), which are supported by a substantially smaller set of parameters (and many of these parameters are physically meaningful) compared to those found in ML models. This dissertation, while apply ML to improve hydrologic modeling, consistently strives to implement ML models with the smallest possible number of parameters, aiming to strike a balance between simplicity and performance. ML models with a larger number of parameters (an extremely complex example is the Large Language Models, LLMs), have demonstrated the potential for superior performance by capturing more

nuanced patterns in data. However, the increased parameter size introduces some challenges, including a heightened risk of overfitting as the model's dimensionality expands, a surge in computational demands for memory storage, and a decline in model interpretability. Model interpretability, in particular, is essential for scientific ML, as it ensures that the outcomes of models can be understood and trusted by the community. Closely tied to the model interpretability, the size of parameters holds similar significance in both ML and process-based models. In hydrologic modeling and similar scientific domains, parameters typically serve to represent aspects of the system that are either not fully comprehended or too intricate for direct inclusion in the model. These parameters act as proxies for specific processes or characteristics that, due to limited knowledge or computational power, cannot be directly observed or accurately modeled. A model with numerous parameters may reflect a lower understanding of physical process, implying reliance on parameter adjustment to align with observed data. Nonetheless, an abundance of parameters is not inherently detrimental; it enables the model to more adaptively represent complex systems. The key challenge lies in balancing model complexity with interpretability and the available calibration data to avoid overfitting. Overfitting occurs when a model, overly dependent on parameters, accurately predicts training data but fails with new data. The goal, therefore, is to strike a balance that ensures the model is both accurate and reliable for its intended purpose, while also being simple enough to facilitate understanding and prevent overfitting.

A feasible way to improve the model interpretability is forcibly "encoding" prior physical knowledge in ML models. This approach can be deemed as substituting a portion of the model's parameters with explicitly represented physical processes, thereby reducing the parameter uncertainty of ML models and avoiding the ill-posedness of the optimization (Rao *et al.*, 2023). In hydrology, one example is the entire hydrological model being hard-coded into a Python version for deployment on a Deep Learning (DL) platform (Tsai *et al.*, 2021), referred as differentiable parameter learning (dPL). It could be seen as an

extreme form of embedding prior physics of the system into ML models by completely substituting ML parameters with the formulas of physical processes. The objective of Tsai *et al.* (2021) is to tackle the challenges associated with parameter estimation and regionalization in hydrologic modeling. Therefore, they maintain limited flexibility for parameters, opting instead to translate the process-based model from Fortran/C++ to Python. In another case, the differentiable, learnable process-based models, referred to as the δ model (Feng *et al.*, 2023), explicitly defined physical processes with flexible modules, which can be optionally replaced by neural networks. The δ model demonstrates their robustness even in the data-scarce regions, which underscores the importance of prior physics on the ML models' interpretability and generalization ability. Looking beyond, there exists the potential to develop a "tightly coupled" scientific knowledge-aware or physics-guided ML model to address hydrologic problems. This is a further step beyond "loosely coupled" physics-guided ML, which involves direct information exchange between the outputs or inputs of the process-based model and the ML model. In a tightly coupled model, operator-level integration could enhance the modeling of physical phenomena beyond the capabilities of traditional process-based approaches. This integration is achieved through the use of automatic differentiation and the explicit encoding of physical equations as part of the loss function or network structure. For instance, large-scale hydrological models and LSMs typically operate at a 1-D column level and often overlook the exchange of horizontal fluxes (Picourlat *et al.*, 2022). However, a tightly coupled physics-guided ML model, which may utilize Convolutional LSTM (ConvLSTM, Shi *et al.*, 2015) or Vision Transformers (ViT, Dosovitskiy *et al.*, 2020) as a foundational framework ("base model") that incorporate physical processes to a certain extent, is capable of capturing complex spatiotemporal dynamics more effectively and reasonably. Specifically, ConvLSTM can be adapted for hydrological modeling by modifying its gating mechanisms or state updates to reflect essential physical laws or constraints, such as the conservation of mass or energy,

ensuring the model's predictions adhere to these fundamental principles. Similarly, ViT could be implemented by embedding physical processes into their self-attention mechanisms, allowing the model to prioritize inputs based on their physical significance, or by incorporating physical constraints directly into the positional encoding to maintain spatial consistency with known physical laws. Such models, once trained, validated, and tested against observational data, may promise to rectify the neglect of horizontal flux exchange seen in traditional models and open up new possibilities for accurately modeling complex hydrological processes. In water management models, incorporating prior physical knowledge ensures that the results are straightforwardly rational. Chapter 3 of the dissertation demonstrates this by applying physical constraints as penalty terms within the loss functions, thereby enhancing the reasonableness of storage simulations (within the operational range). This approach exemplifies the application of physics-guided machine learning (ML) models in water management. For future work on the data-driven reservoir operation models, specific operational rules could be explicitly encoded within the ML models. This, combined with a physics-constrained loss function, will further ensure that the model simulations closely mirror real-world scenarios.

5.2.2 Utilizing Purely ML-based Models for Operational Purposes

Purely ML-based models can also be highly beneficial for operational purposes owing to their high efficiency. As highlighted previously, ML models are inherently data-driven, lacking explicit process representation or a physical basis, which may result in erroneous and unreliable extrapolations. However, ML-based surrogate models could benefit from being informed by the principles underlying physical models. A well-trained and validated surrogate can thus be considered to possess the inherent advantages of process-based models, capable of naturally accommodating spatiotemporal changes. Such surrogates yield predictions that, while not exact, are reasonable and to some extent grounded in physical

understanding (Tran *et al.*, 2023). Therefore, the surrogate can be seen as a valuable tool that partially combines the advantages of both process-based models (physically reasonable) and ML models (fast), albeit not perfectly. As demonstrated in Chapter 2, the computational speed of the ML models can provide a large number of simulations in a short time, facilitating more iterative and exploratory efforts based on a well-validated surrogate model. This capability is valuable in practical management, especially when a comprehensive representation of potential outcomes is required for decision-making, or extensive ensemble simulations are necessary to derive conclusive results that encompass the full range of probabilities. Taking a further development of Chapter 2 as an example, the application of purely ML-based validated surrogate models enables the rapid generation of diverse streamflow predictions across a multitude of vegetation disturbance scenarios (e.g., the unpredictable nature of wildfires, or the strategic implementation of various thinning practices). To some extent, it resembles a form of sensitivity analysis, wherein the effects of varying disturbance characteristics (e.g., extent and severity of wildfire, type of thinning) on streamflow can be explored in detail, which can thereby inform more resilient and adaptive management strategies by understanding how different disturbance scenarios could impact water resources. Additionally, given the rapid simulation capabilities, it is feasible to not rely on one specific surrogate model but instead develop a suite of ML-based surrogates that target outputs from various process-based hydrologic models, which allows for the consideration of uncertainties from different hydrologic model configurations. Similarly, streamflow ensemble forecasting can also be conducted using the validated surrogate of one or more hydrologic models. The typical procedure for producing dynamical streamflow forecasts involves: (a) running a hydrological model with historical data up to the forecast start date to estimate the current hydrologic conditions of the basin, with observational data potentially used to update and correct the model's state variables; and (b) using these hydrologic state estimates to initialize forecast simulations driven by weather forecasts,

thereby producing streamflow forecasts. Weather forecasts may serve as inputs to one or more hydrological models, with the resulting hydrological forecasts being either single-valued or ensemble-based (Troin *et al.*, 2021). The primary motivation for generating ensemble forecasts is to quantify forecast uncertainty. Ensemble Prediction Systems (EPS) have great potential for optimizing prediction confidence levels for risk management (Arduino *et al.*, 2005; Davolio *et al.*, 2013) by sampling uncertainty related to the initial conditions of the atmosphere, Numerical Weather Prediction (NWP), and/or hydrological model formulations. A surrogate-based ensemble can efficiently run under thousands of different initial conditions (initial conditions of the atmosphere and the corresponding streamflow conditions) to mitigate such uncertainties.

REFERENCES

Aboutalebi, M., O. Bozorg Haddad and H. A. Loáiciga, "Optimal monthly reservoir operation rules for hydropower generation derived with SVR-NSGAII", Journal of Water Resources Planning and Management **141**, 11, 04015029 (2015).

Amenu, G. and P. Kumar, "A model for hydraulic redistribution incorporating coupled soil-root moisture transport", Hydrology and Earth System Sciences **12**, 1, 55–74 (2008).

Anderson, J. T. and B.-H. Song, "Plant adaptation to climate change—where are we?", Journal of Systematics and Evolution **58**, 5, 533–545 (2020).

Arduino, G., P. Reggiani and E. Todini, "Recent advances in flood forecasting and flood risk assessment", Hydrology and Earth System Sciences **9**, 4, 280–284 (2005).

Asher, M. J., B. F. Croke, A. J. Jakeman and L. J. Peeters, "A review of surrogate models and their application to groundwater modeling", Water Resources Research **51**, 8, 5957–5973 (2015).

Ashfaq, M., D. Rastogi, J. Kitson, M. Abid and S. Kao, "Evaluation of cmip6 gcms over the conus for downscaling studies", Journal of Geophysical Research: Atmospheres **127**, 21, 2022 036659 (2022).

Babaeian, E., S. Paheding, N. Siddique, V. K. Devabhaktuni and M. Tuller, "Short-and mid-term forecasts of actual evapotranspiration with deep learning", Journal of Hydrology **612**, 128078 (2022).

Barron-Gafford, G. A., E. P. Sanchez-Cañete, R. L. Minor, S. M. Hendryx, E. Lee, L. F. Sutter, N. Tran, E. Parra, T. Colella, P. C. Murphy, E. Hamerlynck, P. Kumar and R. L. Scott, "Impacts of hydraulic redistribution on grass–tree competition vs facilitation in a semi-arid savanna", New Phytologist **215**, 4, 1451–1461 (2017).

Biemans, H., I. Haddeland, P. Kabat, F. Ludwig, R. W. Hutjes, J. Heinke, W. von Bloh and D. Gerten, "Impact of reservoirs on river discharge and irrigation water supply during the 20th century", Water Resources Research **47**, 3 (2011).

Blair, P. and W. Buytaert, "Socio-hydrological modelling: a review asking 'why, what and how?"', Hydrology and Earth System Sciences **20**, 1, 443–478 (2016).

Bollen, K. and J. Pearl, "Eight myths about causality and structural equation models", in "Handbook of causal analysis for social research", p. 301–328 (Springer Netherlands, Dordrecht, 2013).

Bonan, G. B., "Forests and climate change: forcings, feedbacks, and the climate benefits of forests", science **320**, 5882, 1444–1449 (2008).

Boulange, J., N. Hanasaki, D. Yamazaki and Y. Pokhrel, "Role of dams in reducing global flood exposure under climate change", Nature communications **12**, 1, 1–7 (2021).

Boulton, C. A., T. M. Lenton and N. Boers, "Pronounced loss of amazon rainforest resilience since the early 2000s", Nature Climate Change **12**, 3, 271–278 (2022).

Boyle, D. P., H. V. Gupta and S. Sorooshian, "Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods", Water Resources Research **36**, 12, 3663–3674 (2000).

Breiman, L., "Random forests", Machine learning **45**, 5–32 (2001).

Brekke, L., E. Maurer, J. Anderson, M. Dettinger, E. Townsley, A. Harrison and T. Pruitt, "Assessing reservoir operations risk under climate change", Water Resources Research **45**, 4 (2009).

Brown, C. M., J. R. Lund, X. Cai, P. M. Reed, E. A. Zagona, A. Ostfeld, J. Hall, G. W. Characklis, W. Yu and L. Brekke, "The future of water resources systems analysis: Toward a scientific framework for sustainable water management", Water resources research **51**, 8, 6110–6124 (2015).

Broxton, P., X. Zeng and D. N, "Daily 4 km gridded swe and snow depth from assimilated in-situ and modeled data over the conterminous us, version 1", URL https://nsidc.org/data/NSIDC-0719/versions/1., date accessed 01-24-2023. (2019).

Brunner, M. and P. Naveau, "Spatial variability in alpine reservoir regulation: deriving reservoir operations from streamflow using generalized additive models", Hydrology and Earth System Sciences **27**, 3, 673–687 (2023).

Bureau of Reclamation, "2022 annual operating plan for Colorado River reservoirs", Bureau of Reclamation (2022).

Butler, D., "Earth observation enters next phase", Nature **508**, 7495, 160–161 (2014).

Caretta, M. A., A. Mukherji, M. Arfanuzzaman, R. A. Betts, A. Gelfan, Y. Hirabayashi, T. K. Lissner, J. Liu, E. L. Gunn, R. Morgan, S. Mwanga and S. Supratid, *Water*, pp. 551–712 (Cambridge University Press, Cambridge, UK and New York, NY, USA, 2022).

Carriere, P. and R. Wurbs, *Evaluation of storage reallocation and related strategies for optimizing reservoir system operations* (Texas Water Resources Institute, 1988).

Carvalho, D., S. Rafael, A. Monteiro, V. Rodrigues, M. Lopes and A. Rocha, "How well have cmip3, cmip5 and cmip6 future climate projections portrayed the recently observed warming", Scientific Reports **12**, 1, 11983 (2022).

Castle, S., B. Thomas, J. Reager, M. Rodell, S. Swenson and J. Famiglietti, "Groundwater depletion during drought threatens future water security of the Colorado River Basin", Geophysical research letters **41**, 16, 5904–5911 (2014).

Chase-Israel, J., "Building new hydro at an existing water supply lake in Texas", URL https://www.hydro.org/powerhouse/article/building-new-hydro-at-an-existing-water-supply-lake-in-texas/ (2022).

Chen, C., T. Park, X. Wang, S. Piao, B. Xu, R. K. Chaturvedi, R. Fuchs, V. Brovkin, P. Ciais, R. Fensholt, H. Tømmervik, G. Bala, Z. Zhu, R. R. Nemani and R. B. Myneni, "China and India lead in greening of the world through land-use management", Nature sustainability **2**, 2, 122–129 (2019a).

Chen, F. and J. Dudhia, "Coupling an advanced land surface–hydrology model with the penn state–NCAR MM5 modeling system. Part I: Model implementation and sensitivity", Monthly weather review **129**, 4, 569–585 (2001).

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q. Y. Duan, M. Ek and A. Betts, "Modeling of land surface evaporation by four schemes and comparison with FIFE observations", Journal of Geophysical Research: Atmospheres **101**, D3, 7251–7268 (1996).

Chen, J., W. Ju, P. Ciais, N. Viovy, R. Liu, Y. Liu and X. Lu, "Vegetation structural change since 1981 significantly enhanced the terrestrial carbon sink", Nature communications **10**, 1, 4259 (2019b).

Chen, L. and L. Wang, "Recent advance in Earth observation big data for hydrology", Big Earth Data **2**, 1, 86–107 (2018).

Chen, T., C. Song, P. Zhan, J. Yao, Y. Li and J. Zhu, "Remote sensing estimation of the flood storage capacity of basin-scale lakes and reservoirs at high spatial and temporal resolutions", Science of The Total Environment **807**, 150772 (2022a).

Chen, W. and J. Olden, "Designing flows to resolve human and environmental water needs in a dam-regulated river", Nature communications **8**, 1, 1–10 (2017).

Chen, Y., D. Li, Q. Zhao and X. Cai, "Developing a generic data-driven reservoir operation model", Advances in Water Resources **167**, 104274 (2022b).

Cheng, L., M. Hoerling, L. Smith and J. Eischeid, "Diagnosing human-induced dynamic and thermodynamic drivers of extreme rainfall", Journal of Climate **31**, 3, 1029–1051 (2018).

Christensen, N. and D. Lettenmaier, "A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River Basin", Hydrology and Earth System Sciences **11**, 4, 1417–1434 (2007).

Christensen, N., A. Wood, N. Voisin, D. Lettenmaier and R. Palmer, "The effects of climate change on the hydrology and water resources of the Colorado River Basin", Climatic change **62**, 337–363 (2004).

Cihlar, J., J. Chen and Z. Li, "On the validation of satellite-derived products for land applications", Canadian journal of remote sensing **23**, 4, 381–389 (1997).

Clapp, R. B. and G. M. Hornberger, "Empirical equations for some soil hydraulic properties", Water resources research **14**, 4, 601–604 (1978).

Clark, M., Y. Fan, D. Lawrence, J. Adam, D. Bolster, D. Gochis and X. Zeng, "Improving the representation of hydrologic processes in Earth system models", Water Resources Research **51**, 8, 5929–5956 (2015).

Clark, M. P., M. F. Bierkens, L. Samaniego, R. A. Woods, R. Uijlenhoet, K. E. Bennett, V. Pauwels, X. Cai, A. W. Wood and C. D. Peters-Lidard, "The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism", Hydrology and Earth System Sciences **21**, 7, 3427–3440 (2017).

Clark, M. P., D. Kavetski and F. Fenicia, "Pursuing the method of multiple working hypotheses for hydrological modeling", Water Resources Research **47**, 9 (2011).

Claverie, M., J. Matthews, E. Vermote and C. Justice, "A 30+ year avhrr lai and fapar climate data record: Algorithm description and validation", Remote Sensing **8**, 3, 263 (2016).

Coenders-Gerrits, A., R. Ent, T. Bogaard, L. Wang-Erlandsson, M. Hrachowitz and H. Savenije, "Uncertainties in transpiration estimates", Nature **506**, 7487, 1– 2 (2014).

Coerver, H., M. Rutten and N. Giesen, "Deduction of reservoir operating rules for application in global hydrological models", Hydrology and Earth System Sciences **22**, 1, 831–851 (2018).

Cohen, M., J. Christian-Smith and J. Berggren, *Water to Supply the Land: Irrigated Agriculture in the Colorado River Basin* (Pacific Institute, Oakland, CA, 2013).

Condon, L., R. Maxwell, V. Bansal, L. Bearup, C. Chennault, L. Gallagher, S. Gangopadhyay, R. Hull, T. Kelly, C. Kreitzberg, E. Leonarduzzi, P. Melchior, N. Merchant, E. Skidmore, T. Swetnam, S. Spraragen, H. Tran and J. Williams, "Lessons learned designing the hydrogen machine learning platform for hydrologic exploration: a story of collaboration between hydrologic scientists, software developers, machine learning researchers and water managers", in "AGU Fall Meeting Abstracts", vol. 2021, pp. SY24A–04 (2021).

Condon, L. E. and R. M. Maxwell, "Simulating the sensitivity of evapotranspiration and streamflow to large-scale groundwater depletion", Science Advances **5**, 6, eaav4574 (2019).

Cosgrove, W. and D. Loucks, "Water management: Current and future challenges and research directions", Water Resources Research **51**, 6, 4823–4839 (2015).

Croley, T. and K. Raja Rao, "Multiobjective risks in reservoir operation", Water Resources Research **15**, 4, 807–814 (1979).

Currier, W., A. Wood, N. Mizukami, B. Nijssen, J. Hamman and E. Gutmann, "Vegetation representation influences projected streamflow changes in the Colorado River Basin", Journal of Hydrometeorology (2023).

Dannenberg, M., E. Wise and W. Smith, "Reduced tree growth in the semiarid United States due to asymmetric responses to intensifying precipitation extremes", Science advances **5**, 10, 0667 (2019).

Davolio, S., M. Miglietta, T. Diomede, C. Marsigli and A. Montani, "A flood episode in northern italy: multi-model and single-model mesoscale meteorological ensembles for hydrological predictions", Hydrology and Earth System Sciences **17**, 6, 2107–2120 (2013).

Delaney, C., R. Hartman, J. Mendoza, M. Dettinger, L. Delle Monache, J. Jasperse and S. Evett, "Forecast informed reservoir operations using ensemble streamflow predictions for a multipurpose reservoir in northern california", Water Resources Research **56**, 9, 2019 026604 (2020).

Denager, T., T. Sonnenborg, M. Looms, H. Bogena and K. Jensen, "Point-scale multi-objective calibration of the Community Land Model (version 5.0) using in situ observations of water and energy fluxes and variables", Hydrology and Earth System Sciences **27**, 14, 2827–2845 (2023).

Denaro, S., D. Anghileri, M. Giuliani and A. Castelletti, "Informing the operations of water reservoirs over multiple temporal scales by direct use of hydro-meteorological data", Advances in water resources **103**, 51–63 (2017).

DeNeale, S. T., G. B. Baecher, K. M. Stewart, E. D. Smith and D. B. Watson, "Current state-of-practice in dam safety risk assessment", Tech. rep., Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States) (2019).

Ding, H. and Y. Zhu, "NDE vegetation products system algorithm theoretical basis document, version 4.0", URL https://www.ospo.noaa.gov/Products/documents/GVF_ATBD_V4.0.pdf, nOAA/NESDIS/OSPO (2018).

Ding, W., C. Zhang, Y. Peng, R. Zeng, H. Zhou and X. Cai, "An analytical framework for flood water conservation considering forecast uncertainty and acceptable risk", Water Resources Research **51**, 6, 4702–4726 (2015).

Döll, P., K. Fiedler and J. Zhang, "Global-scale analysis of river flow alterations due to water withdrawals and reservoirs", Hydrology and Earth System Sciences **13**, 12, 2413–2432 (2009).

Döll, P. and H. Schmied, "How is the impact of climate change on river flow regimes related to the impact on mean annual runoff? a global-scale analysis", Environmental Research Letters **7**, 1, 014037 (2012).

Dong, N., W. Guan, J. Cao, Y. Zou, M. Yang, J. Wei, L. Chen and H. Wang, "A hybrid hydrologic modelling framework with data-driven and conceptual reservoir operation schemes for reservoir impact assessment and predictions", Journal of Hydrology **129246** (2023).

Dosovitskiy, A., L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale", arXiv preprint arXiv:2010.11929 (2020).

Doyle, M. and L. Patterson, "Federal decentralization and adaptive management of water resources: reservoir reallocation by the us army corps of engineers", JAWRA Journal of the American Water Resources Association **55**, 5, 1248–1267 (2019).

Duarte, C., I. Losada, I. Hendriks, I. Mazarrasa and N. Marbà, "The role of coastal plant communities for climate change mitigation and adaptation", Nature climate change **3**, 11, 961–968 (2013).

Duveiller, G., J. Hooker and A. Cescatti, "The mark of vegetation change on Earth's surface energy balance", Nature communications **9**, 1, 1–12 (2018).

Ehsani, N., C. J. Vörösmarty, B. M. Fekete and E. Z. Stakhiv, "Reservoir operations under climate change: Storage capacity options to mitigate risk", Journal of Hydrology **555**, 435–446 (2017).

Eilander, D., F. Annor, L. Iannini and N. Giesen, "Remotely sensed monitoring of small reservoir dynamics: A bayesian approach", Remote Sensing **6**, 2, 1191–1210 (2014).

Ek, M., K. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno and J. Tarpley, "Implementation of noah land surface model advances in the national centers for environmental prediction operational mesoscale eta model", Journal of Geophysical Research: Atmospheres **108**, D22 (2003).

Elmer, N., J. Mecikalski, A. Molthan, D. Gochis and U. Nair, "Impact of real-time satellite-derived green vegetation fraction on national water model fluxes in humid alabama watersheds", JAWRA Journal of the American Water Resources Association **58**, 1, 104–118 (2022).

Energy Information Administration, "Today in energy - specific article title", URL https://www.eia.gov/todayinenergy/detail.php?id=55960 (2022).

Feng, D., H. Beck, K. Lawson and C. Shen, "The suitability of differentiable, physics-informed machine learning hydrologic models for ungauged regions and climate change impact assessment", Hydrology and Earth System Sciences **27**, 12 (2023).

Feng, D., K. Fang and C. Shen, "Enhancing streamflow forecast and extracting insights using long-short term memory networks with data integration at continental scales", Water Resources Research **56**, 9, 2019 026793 (2020).

Ferguson, I. M. and R. M. Maxwell, "Hydrologic and land–energy feedbacks of agricultural water management practices", Environmental Research Letters **6**, 1, 014006 (2011).

Fernando, D., Y. Zhang, S. Fakhreddine and B. Scanlon, "Forecast-informed reservoir operations in Texas: what will it take to say 'yes'?", in "AGU Fall Meeting Abstracts", vol. 2020, p. 015–0008 (2020).

Ficklin, D., I. Stewart and E. Maurer, "Climate change impacts on streamflow and subbasin-scale hydrology in the Upper Colorado River Basin", PloS one **8**, 8, 71297 (2013).

Fisher, A., C. Rudin and F. Dominici, "All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously.", J. Mach. Learn. Res. **20**, 177, 1–81 (2019).

Forsberg, B., J. Melack, T. Dunne, R. Barthem, M. Goulding, R. Paiva and S. Weisser, "The potential impact of new andean dams on amazon fluvial ecosystems", PloS one **12**, 8, 0182254 (2017).

Friedrich, K., R. Grossman, J. Huntington, P. Blanken, J. Lenters, K. Holman and T. Kowalski, "Reservoir evaporation in the western United States: current science, challenges, and future needs", Bulletin of the American Meteorological Society **99**, 1, 167–187 (2018).

Gangrade, S., D. Lu, S. Kao and S. Painter, "Machine learning assisted reservoir operation model for long-term water management simulation", JAWRA Journal of the American Water Resources Association (2022).

George, E., W. Horst and E. Neumann, "Adaptation of plants to adverse chemical soil conditions", in "Marschner's mineral nutrition of higher plants", p. 409–472 (Academic press, 2012).

Gettelman, A., A. J. Geer, R. M. Forbes, G. R. Carmichael, G. Feingold, D. J. Posselt, G. L. Stephens, S. C. van den Heever, A. C. Varble and P. Zuidema, "The future of Earth system prediction: Advances in model-data fusion", Science Advances **8**, 14, eabn3488 (2022).

Giuliani, M., J. Lamontagne, P. Reed and A. Castelletti, "A state-of-the-art review of optimal reservoir control for managing conflicting demands in a changing world", Water Resources Research **57**, 12, 2021 029927 (2021).

Gochis, D., M. Barlage, A. Dugger, K. FitzGerald, L. Karsten, M. McAllister, J. McCreight, J. Mills, A. RafieeiNasab, L. Read and K. Sampson, "The wrf-hydro modeling system technical description", (2018).

Good, S., D. Noone and G. Bowen, "Hydrologic connectivity constrains partitioning of global terrestrial water fluxes", Science **349**, 6244, 175–177 (2015).

Graf, W., "Dam nation: A geographic census of american dams and their large-scale hydrologic impacts", Water resources research **35**, 4, 1305–1311 (1999).

Grossiord, C., C. Bachofen, J. Gisler, E. Mas, Y. Vitasse and M. Didion-Gency, "Warming may extend tree growing seasons and compensate for reduced carbon uptake during dry periods", Journal of Ecology **110**, 7, 1575–1589 (2022).

Gutman, G. and A. Ignatov, "The derivation of the green vegetation fraction from noaa/avhrr data for use in numerical weather prediction models", International Journal of remote sensing **19**, 8, 1533–1543 (1998).

Haddeland, I., T. Skaugen and D. P. Lettenmaier, "Anthropogenic impacts on continental surface water fluxes", Geophysical Research Letters **33**, 8 (2006).

Hadjerioua, B., Y. Wei and S. Kao, "An assessment of energy potential at non-powered dams in the United States", Tech. rep., EERE Publication and Product Library, Washington, DC (United States) (2012).

Hadjimichael, A., J. Yoon, P. Reed, N. Voisin and W. Xu, "Exploring the consistency of water scarcity inferences between large-scale hydrologic and node-based water system model representations of the Upper Colorado River Basin", Journal of Water Resources Planning and Management **149**, 2, 04022081 (2023).

Hamman, J., B. Nijssen, T. Bohn, D. Gergel and Y. Mao, "The variable infiltration capacity model version 5 (vic-5): Infrastructure improvements for new applications and reproducibility", Geoscientific Model Development **11**, 8, 3481–3496 (2018).

Harding, B., A. Wood and J. Prairie, "The implications of climate change scenario selection for future streamflow projection in the Upper Colorado River Basin", Hydrology and Earth System Sciences **16**, 11, 3989–4007 (2012).

Hashimoto, T., J. Stedinger and D. Loucks, "Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation", Water resources research **18**, 1, 14–20 (1982).

Hejazi, M. I., X. Cai and B. L. Ruddell, "The role of hydrologic information in reservoir operation–learning from historical releases", Advances in water resources **31**, 12, 1636–1650 (2008).

Hibbert, A., *Managing vegetation to increase flow in the Colorado River Basin*, vol. 66 (Rocky Mountain Forest and Range Experiment Station, Forest Service, 1979).

Hidalgo, H., T. Piechota and J. Dracup, "Alternative principal components regression procedures for dendrohydrologic reconstructions", Water Resources Research **36**, 11, 3241–3249 (2000).

Hipni, A., A. El-shafie, A. Najah, O. Karim, A. Hussain and M. Mukhlisin, "Daily forecasting of dam water levels: comparing a support vector machine (svm) model with adaptive neuro fuzzy inference system (anfis", Water resources management **27**, 10, 3803–3823 (2013).

Ho, M., U. Lall, M. Allaire, N. Devineni, H. Kwon, I. Pal and D. Wegner, "The future role of dams in the United States of america", Water Resources Research **53**, 2, 982–998 (2017).

Hochreiter, S. and J. Schmidhuber, "Long short-term memory", Neural computation **9**, 8, 1735–1780 (1997).

Hoerling, M., A. Kumar, R. Dole, J. Nielsen-Gammon, J. Eischeid, J. Perlwitz, X. Quan, T. Zhang, P. Pegion and M. Chen, "Anatomy of an extreme event", Journal of Climate **26**, 9, 2811–2832 (2013).

Iglesias, A. and L. Garrote, "Adaptation strategies for agricultural water management under climate change in europe", Agricultural water management **155**, 113–124 (2015).

Jacobs, J., "The sustainability of water resources in the Colorado River Basin", Bridge **41**, 4, 6–12 (2011).

Jain, S., G. Yoganarasimhan and S. Seth, "A risk-based approach for flood control operation of a multipurpose reservoir", Jawra Journal of the American Water Resources Association **28**, 6, 1037–1043 (1992).

Jakeman, A. J., R. A. Letcher and J. P. Norton, "Ten iterative steps in development and evaluation of environmental models", Environmental Modelling & Software **21**, 5, 602–614 (2006).

Jasperse, J., F. Ralph, M. Anderson, L. Brekke, N. Malasavage, M. Dettinger, J. Forbis, J. Fuller, C. Talbot, R. Webb and A. Haynes, "Lake Mendocino forecast informed reservoir operations final viability assessment", (2020).

Jia, W., Y. Zhang, Z. Wei, Z. Zheng and P. Xie, "Daily reference evapotranspiration prediction for irrigation scheduling decisions based on the hybrid pso-lstm model", Plos one **18**, 4, 0281478 (2023).

Jiang, Z., M. Vargas and I. Csiszar, "New oprational real-time daily rolling weekly green vegetation fraction product derived from suomi npp viirs reflectance data", in "2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS", p. 3524–3527 (IEEE, 2016).

Kao, S., M. Ashfaq, D. Rastogi, S. Gangrade, R. Uria Martinez, A. Fernandez, G. Konapala, N. Voisin, T. Zhou, W. Xu and H. Gao, "The third assessment of the effects of climate change on federal hydropower", Tech. rep., Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States) (2022).

Kao, S.-C., M. Ashfaq, B. S. Naz, R. Uria Martinez, D. Rastogi, R. Mei, Y. Jager, N. M. Samu and M. J. Sale, "The second assessment of the effects of climate change on federal hydropower", Tech. rep., Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States) (2016).

Karsanina, M. and K. Gerke, "Hierarchical optimization: Fast and robust multiscale stochastic reconstructions with rescaled correlation functions", Physical review letters **121**, 26, 265501 (2018).

Kasim, M., D. Watson-Parris, L. Deaconu, S. Oliver, P. Hatfield, D. Froula, G. Gregori, M. Jarvis, S. Khatiwala, J. Korenaga, J. Topp-Mugglestone, E. Viezzer and S. Vinko, "Building high accuracy emulators for scientific simulations with deep neural architecture search", Machine Learning: Science and Technology **3**, 1, 015013 (2021).

Kawashima, S., "Conserving reservoir water storage: An economic appraisal", Water resources research **43**, 5 (2007).

Kelly, A. E. and M. L. Goulden, "Rapid shifts in plant distribution with recent climate change", Proceedings of the national academy of sciences **105**, 33, 11823–11826 (2008).

Khazaei, B., L. K. Read, M. Casali, K. M. Sampson and D. Yates, "Improvement of lake and reservoir parameterization in the noaa national water model", in "World Environmental and Water Resources Congress 2021", pp. 552–560 (2021).

Kim, S., H. Sadeghi, R. Limon, M. Saharia, D. Seo, A. Philpott, F. Bell, J. Brown and M. He, "Assessing the skill of medium-range ensemble precipitation and streamflow forecasts from the hydrologic ensemble forecast service (HEFS) for the upper trinity river basin in north Texas", Journal of Hydrometeorology **19**, 9, 1467–1483 (2018).

Kingma, D. P. and J. Ba, "Adam: A method for stochastic optimization", arXiv preprint arXiv:1412.6980 (2014).

Kline, R. B., *Principles and practice of structural equation modeling* (Guilford publications, 2023).

Klipsch, J. and M. Hurst, *HEC-ResSim reservoir system simulation user's manual version 3.0* (USACE, Davis, CA, 2007).

Kolassa, J., R. H. Reichle, R. D. Koster, Q. Liu, S. Mahanama and F.-W. Zeng, "An observation-driven approach to improve vegetation phenology in a global land surface model", Journal of Advances in Modeling Earth Systems **12**, 9, e2020MS002083 (2020).

Konapala, G., A. Mishra, Y. Wada and M. Mann, "Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation", Nature communications **11**, 1, 3044 (2020).

Konieczki, M., C. Ottman, C. Jalbert, A. Godbey and A. Ickert, "Forecast-informed reservoir operations: A case study of the west fork of the trinity river", in "World Environmental and Water Resources Congress 2017", pp. 584–594 (2017).

Konings, A. G. and P. Gentine, "Global variations in ecosystem-scale isohydricity", Global change biology **23**, 2, 891–905 (2017).

Koster, R. D., G. K. Walker, G. J. Collatz and P. E. Thornton, "Hydroclimatic controls on the means and variability of vegetation phenology and carbon uptake", Journal of climate **27**, 14, 5632–5652 (2014).

Kratzert, F., D. Klotz, C. Brenner, K. Schulz and M. Herrnegger, "Rainfall–runoff modelling using long short-term memory (lstm) networks", Hydrology and Earth System Sciences **22**, 11, 6005–6022 (2018).

Kratzert, F., D. Klotz, M. Herrnegger, A. Sampson, S. Hochreiter and G. Nearing, "Toward improved predictions in ungauged basins: Exploiting the power of machine learning", Water Resources Research **55**, 12, 11344–11354 (2019).

Krzysztofowicz, R. and L. Duckstein, "Preference criterion for flood control under uncertainty", Water Resources Research **15**, 3, 513–520 (1979).

Kumar, S. V., D. M. Mocko, S. Wang, C. D. Peters-Lidard and J. Borak, "Assimilation of remotely sensed leaf area index into the Noah-MP land surface model: Impacts on water and carbon fluxes and states over the continental United States", Journal of Hydrometeorology **20**, 7, 1359–1377 (2019).

Kumar, S. V., C. D. Peters-Lidard, Y. Tian, P. R. Houser, J. Geiger, S. Olden, L. Lighty, J. L. Eastman, B. Doty, P. Dirmeyer and J. Adams, "Land information system: An interoperable framework for high resolution land surface modeling", Environmental modelling & software **21**, 10, 1402–1415 (2006).

Kumar, S. V., S. Wang, D. M. Mocko, C. D. Peters-Lidard and Y. Xia, "Similarity assessment of land surface model outputs in the north american land data assimilation system", Water Resources Research **53**, 11, 8941–8965 (2017).

Kunkel, K. E., D. R. Easterling, K. Hubbard and K. Redmond, "Temporal variations in frost-free season in the United States: 1895–2000", Geophysical Research Letters **31**, 3 (2004).

Lane, N., *Aging infrastructure: Dam safety* (Congressional Research Service, 2007).

Le Page, M., Y. Fakir, L. Jarlan, A. Boone, B. Berjamy, S. Khabba and M. Zribi, "Projection of irrigation water demand based on the simulation of synthetic crop coefficients and climate change", Hydrology and Earth System Sciences **25**, 2, 637–651 (2021).

Lee, E., F.-W. Zeng, R. D. Koster, B. Weir, L. E. Ott and B. Poulter, "The impact of spatiotemporal variability in atmospheric CO_2 concentration on global terrestrial carbon fluxes", Biogeosciences **15**, 18, 5635–5652 (2018).

Lehner, B., C. Liermann, C. Revenga, C. Vörösmarty, B. Fekete, P. Crouzet and D. Wisser, "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 (2011).

Li, C., G. Sun, P. V. Caldwell, E. Cohen, Y. Fang, Y. Zhang, L. Oudin, G. M. Sanchez and R. K. Meentemeyer, "Impacts of urbanization on watershed water balances across the conterminous United States", Water Resources Research **56**, 7, e2019WR026574 (2020).

Li, D., M. L. Wrzesien, M. Durand, J. Adam and D. P. Lettenmaier, "How much runoff originates as snow in the western United States, and how will that change in the future?", Geophysical Research Letters **44**, 12, 6163–6172 (2017).

Li, Q., Y. Zhu, W. Shangguan, X. Wang, L. Li and F. Yu, "An attention-aware LSTM model for soil moisture and soil temperature prediction", Geoderma **409**, 115651 (2022).

Lian, X., S. Piao, L. Z. Li, Y. Li, C. Huntingford, P. Ciais, A. Cescatti, I. A. Janssens, J. Peñuelas, W. Buermann and A. Chen, "Summer soil drying exacerbated by earlier spring greening of northern vegetation", Science advances **6**, 1, eaax0255 (2020).

Liang, X., D. P. Lettenmaier, E. F. Wood and S. J. Burges, "A simple hydrologically based model of land surface water and energy fluxes for general circulation models", Journal of Geophysical Research: Atmospheres **99**, D7, 14415–14428 (1994).

Lin, J., C. Cheng and K. Chau, "Using support vector machines for long-term discharge prediction", Hydrological sciences journal **51**, 4, 599–612 (2006).

Link, T. E. and D. Marks, "Point simulation of seasonal snow cover dynamics beneath boreal forest canopies", Journal of Geophysical Research: Atmospheres **104**, D22, 27841–27857 (1999).

Liu, J., T. Dietz, S. Carpenter, C. Folke, M. Alberti, C. Redman and W. Provencher, "Coupled human and natural systems", AMBIO: a journal of the human environment **36**, 8, 639–649 (2007).

Liu, Q., S. Piao, I. A. Janssens, Y. Fu, S. Peng, X. Lian, P. Ciais, R. B. Myneni, J. Peñuelas and T. Wang, "Extension of the growing season increases vegetation exposure to frost", Nature communications **9**, 1, 426 (2018).

Liu, Y. and H. V. Gupta, "Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework", Water resources research **43**, 7 (2007).

Liu, Y., R. Liu and J. M. Chen, "Retrospective retrieval of long-term consistent global leaf area index (1981–2011) from combined AVHRR and MODIS data", Journal of Geophysical Research: Biogeosciences **117**, G4 (2012).

Livneh, B., T. Bohn, D. Pierce, F. Munoz-Arriola, B. Nijssen, R. Vose, D. Cayan and L. Brekke, "A spatially comprehensive, hydrometeorological data set for mexico, the us, and southern canada 1950–2013", Scientific data **2**, 1, 1–12 (2015).

Livneh, B., Y. Xia, K. E. Mitchell, M. B. Ek and D. P. Lettenmaier, "Noah LSM snow model diagnostics and enhancements", Journal of Hydrometeorology **11**, 3, 721–738 (2010).

Longyang, Q. and R. Zeng, "A hierarchical temporal scale framework for data-driven reservoir release modeling", Water Resources Research **59**, 6, e2022WR033922 (2023a).

Longyang, Q. and R. Zeng, "A hierarchical temporal scale framework for data-driven reservoir release modeling", Water Resources Research **59**, 6, e2022WR033922 (2023b).

Lundberg, S. and S. Lee, "A unified approach to interpreting model predictions", Advances in neural information processing systems **30** (2017).

Ma, K., D. Feng, K. Lawson, W.-P. Tsai, C. Liang, X. Huang, A. Sharma and C. Shen, "Transferring hydrologic data across continents–leveraging data-rich regions to improve hydrologic prediction in data-sparse regions", Water Resources Research **57**, 5, e2020WR028600 (2021).

Ma, T., Z. Duan, R. Li and X. Song, "Enhancing SWAT with remotely sensed lai for improved modelling of ecohydrological process in subtropics", Journal of hydrology **570**, 802–815 (2019).

Mahrt, L. and M. Ek, "The influence of atmospheric stability on potential evaporation", Journal of Applied Meteorology and Climatology **23**, 2, 222–234 (1984).

Malle, J., N. Rutter, G. Mazzotti and T. Jonas, "Shading by trees and fractional snow cover control the subcanopy radiation budget", Journal of Geophysical Research: Atmospheres **124**, 6, 3195–3207 (2019).

Martel, J.-L., F. Brissette, M. Troin, R. Arsenault, J. Chen, T. Su and P. Lucas-Picher, "Cmip5 and cmip6 model projection comparison for hydrological impacts over north america", Geophysical Research Letters **49**, 15, e2022GL098364 (2022).

Masson-Delmotte, V., P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis and M. Huang, "Climate change 2021: the physical science basis", Contribution of working group I To the sixth assessment report of the intergovernmental panel on climate change, 2. (2021).

Mateo, C., N. Hanasaki, D. Komori, K. Tanaka, M. Kiguchi, A. Champathong and T. Oki, "Assessing the impacts of reservoir operation to floodplain inundation by combining hydrological, reservoir management, and hydrodynamic models", Water Resources Research **50**, 9, 7245–7266 (2014).

Maxwell, R. M., L. E. Condon and P. Melchior, "A physics-informed, machine learning emulator of a 2D surface water model: what temporal networks and simulation-based inference can help us learn about hydrologic processes", Water **13**, 24, 3633 (2021).

McCabe, G. J. and D. M. Wolock, "The water-year water balance of the Colorado River Basin", JAWRA Journal of the American Water Resources Association **56**, 4, 724–737 (2020).

McMillan, H. K., I. K. Westerberg and T. Krueger, "Hydrological data uncertainty and its implications", Wiley Interdisciplinary Reviews: Water **5**, 6, e1319 (2018).

Meempatta, L., A. J. Webb, A. C. Horne, L. A. Keogh, A. Loch and M. J. Stewardson, "Reviewing the decision-making behavior of irrigators", Wiley Interdisciplinary Reviews: Water **6**, 5, e1366 (2019).

Mekonnen, M. and A. Hoekstra, "The blue water footprint of electricity from hydropower", Hydrology and Earth System Sciences **16**, 1, 179–187 (2012).

Meyal, A. Y., R. Versteeg, E. Alper, D. Johnson, A. Rodzianko, M. Franklin and H. Wainwright, "Automated cloud based long short-term memory neural network based SWE prediction", Frontiers in Water **2**, 574917 (2020).

Miller, J., M. Barlage, X. Zeng, H. Wei, K. Mitchell and D. Tarpley, "Sensitivity of the NCEP/Noah land surface model to the MODIS green vegetation fraction data set", Geophysical research letters **33**, 13 (2006).

Milly, P., J. Betancourt, M. Falkenmark, R. Hirsch, Z. Kundzewicz, D. Lettenmaier and R. Stouffer, "Stationarity is dead: Whither water management?", Science **319**, 5863, 573–574 (2008a).

Milly, P. and K. Dunne, "Potential evapotranspiration and continental drying", Nature Climate Change **6**, 10, 946–949 (2016).

Milly, P. C., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier and R. J. Stouffer, "Stationarity is dead: Whither water management?", Science **319**, 5863, 573–574 (2008b).

Miranda, J. d. D., C. Armas, F. Padilla and F. Pugnaire, "Climatic change and rainfall patterns: effects on semi-arid plant communities of the Iberian southeast", Journal of Arid Environments **75**, 12, 1302–1309 (2011).

Moran, E., M. Lopez, N. Moore, N. Müller and D. Hyndman, "Sustainable hydropower in the 21st century", Proceedings of the National Academy of Sciences **115**, 47, 11891–11898 (2018).

Mostafaie, A., E. Forootan, A. Safari and M. Schumacher, "Comparing multi-objective optimization techniques to calibrate a conceptual hydrological model using in situ runoff and daily GRACE data", Computational Geosciences **22**, 789–814 (2018).

Muller, C. and P. O'Gorman, "An energetic perspective on the regional response of precipitation to climate change", Nature Climate Change **1**, 5, 266–271 (2011).

Murray-Tortarolo, G., A. Anav, P. Friedlingstein, S. Sitch, S. Piao, Z. Zhu, B. Poulter, S. Zaehle, A. Ahlström, M. Lomas, S. Levis, N. Viovy and N. Zeng, "Evaluation of land surface models in reproducing satellite-derived LAI over the high-latitude northern hemisphere. Part I: Uncoupled DGVMs", Remote Sensing **5**, 10, 4819–4838 (2013).

Nash, J. and J. Sutcliffe, "River flow forecasting through conceptual models part i—a discussion of principles", Journal of hydrology **10**, 3, 282–290 (1970).

National Inventory of Dams, URL https://nid.sec.usace.army.mil/#/, retrieved [Jan 25th, 2024] (n.d.).

Nielsen-Gammon, J., J. Banner, B. Cook, D. Tremaine, C. Wong, R. Mace, H. Gao, Z. Yang, M. Gonzalez, R. Hoffpauir and T. Gooch, "Unprecedented drought challenges for Texas water resources in a changing climate: what do researchers and stakeholders need to know?", Earth's Future **8**, 8, 2020 001552 (2020).

Niu, G., Z. Yang, K. Mitchell, F. Chen, M. Ek, M. Barlage, A. Kumar, K. Manning, D. Niyogi, E. Rosero, M. Tewari and Y. Xia, "The community noah land surface model with multiparameterization options (noah-mp): 1. model description and evaluation with local-scale measurements", Journal of Geophysical Research: Atmospheres **116**, D12 (2011).

Nobre, C., G. Sampaio, L. Borma, J. Castilla-Rubio, J. Silva and M. Cardoso, "Land-use and climate change risks in the amazon and the need of a novel sustainable development paradigm", Proceedings of the National Academy of Sciences **113**, 39, 10759–10768 (2016).

Nogueira, M., S. Boussetta, G. Balsamo, C. Albergel, I. Trigo, F. Johannsen and E. Dutra, "Upgrading land-cover and vegetation seasonality in the ecmwf coupled system: Verification with fluxnet sites, meteosat satellite land surface temperatures, and era5 atmospheric reanalysis", Journal of Geophysical Research: Atmospheres **126**, 15, 2020 034163 (2021).

Noilhan, J. and S. Planton, "A simple parameterization of land surface processes for meteorological models", Monthly weather review **117**, 3, 536–549 (1989).

Norman, J., W. Kustas and K. Humes, "Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature", Agricultural and Forest Meteorology **77**, 3-4, 263–293 (1995).

Oehri, J., G. Schaepman-Strub, J. Kim, R. Grysko, H. Kropp, I. Grünberg, V. Zemlianskii, O. Sonnentag, E. Euskirchen, M. Reji Chacko and G. Muscari, "Vegetation type is an important predictor of the arctic summer land surface energy budget", Nature communications **13**, 1, 6379 (2022).

Opperman, J., R. Camargo, A. Laporte-Bisquit, C. Zarfl and A. Morgan, "Using the wwf water risk filter to screen existing and projected hydropower projects for climate and biodiversity risks", Water **14**, 5, 721 (2022).

Orth, R., "Global soil moisture data derived through machine learning trained with in-situ measurements", Scientific Data **8**, 1, 1–14 (2021).

Ortiz-Partida, J., B. Lane and S. Sandoval-Solis, "Economic effects of a reservoir re-operation policy in the rio grande/bravo for integrated human and environmental water management", Journal of Hydrology: Regional Studies **8**, 130–144 (2016).

Oudin, L., B. Salavati, C. Furusho-Percot, P. Ribstein and M. Saadi, "Hydrological impacts of urbanization at the catchment scale", Journal of Hydrology **559**, 774–786 (2018).

O'Neill, B., C. Tebaldi, D. Vuuren, V. Eyring, P. Friedlingstein, G. Hurtt, R. Knutti, E. Kriegler, J. Lamarque, J. Lowe and G. Meehl, "The scenario model intercomparison project (scenariomip) for cmip6", Geoscientific Model Development **9**, 9, 3461–3482 (2016).

Palmer, M. and A. Ruhi, "Linkages between flow regime, biota, and ecosystem processes: Implications for river restoration", Science **365**, 6459, 2087 (2019).

Pathak, R., H. Dasari, K. Ashok and I. Hoteit, "Effects of multi-observations uncertainty and models similarity on climate change projections", npj Climate and Atmospheric Science **6**, 1, 144 (2023).

Patterson, L. and M. Doyle, "A nationwide analysis of us army corps of engineers reservoir performance in meeting operational targets", JAWRA Journal of the American Water Resources Association **54**, 2, 543–564 (2018).

Picourlat, F., E. Mouche and C. Mügler, "Upscaling hydrological processes for land surface models with a two-hydrologic-variable model: application to the little washita watershed", Water Resources Research **58**, 9, e2021WR030997 (2022).

Pokhrel, Y., F. Felfelani, Y. Satoh, J. Boulange, P. Burek, A. Gädeke, D. Gerten, S. Gosling, M. Grillakis, L. Gudmundsson and N. Hanasaki, "Global terrestrial water storage and drought severity under climate change", Nature Climate Change **11**, 3, 226–233 (2021).

Poulter, B., D. Frank, P. Ciais, R. Myneni, N. Andela, J. Bi and G. Werf, "Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle", Nature **509**, 7502, 600–603 (2014).

Qiu, D., R. Xu, C. Wu, X. Mu, G. Zhao and P. Gao, "Vegetation restoration improves soil hydrological properties by regulating soil physicochemical properties in the loess plateau", China. Journal of Hydrology **609**, 127730 (2022).

Quinn, J., P. Reed, M. Giuliani and A. Castelletti, "What is controlling our control rules? opening the black box of multireservoir operating policies using time-varying sensitivity analysis", Water Resources Research **55**, 7, 5962–5984 (2019).

Ramstein, G., H. Boer and W. Soh, "Vegetation and climate interactions: an introduction", in "EGU General Assembly Conference Abstracts", p. 19655 (2017).

Rao, C., P. Ren, Q. Wang, O. Buyukozturk, H. Sun and Y. Liu, "Encoding physics to learn reaction–diffusion processes", Nature Machine Intelligence **5**, 7, 765–779 (2023).

Rao, I., J. Miles, S. Beebe and W. Horst, "Root adaptations to soils with low fertility and aluminium toxicity", Annals of Botany **118**, 4, 593–605 (2016).

Raso, L., D. Schwanenberg, N. Giesen and P. Overloop, "Short-term optimal operation of water systems using ensemble forecasts", Advances in water resources **71**, 200–208 (2014).

Rastogi, D., S. Kao and M. Ashfaq, "How may the choice of downscaling techniques and meteorological reference observations affect future hydroclimate projections?", Earth's Future **10**, 8, 2022 002734 (2022).

Reichstein, M., G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais and f. Prabhat, "Deep learning and process understanding for data-driven Earth system science", Nature **566**, 7743, 195–204 (2019).

Robbins, H. and S. Monro, "A stochastic approximation method", The annals of mathematical statistics p. 400–407 (1951).

Rocks, J. W. and P. Mehta, "Memorizing without overfitting: Bias, variance, and interpolation in overparameterized models", Physical review research **4**, 1, 013201 (2022).

Ruhge, R. L. and M. Barlage, "Integrating a real-time green vegetation fraction (gvf) product into the land information system (lis)", in "15th Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface", (2011).

Ruiz-Vásquez, M., S. O, G. Arduini, S. Boussetta, A. Brenning, A. Bastos, S. Koirala, G. Balsamo, M. Reichstein and R. Orth, "Impact of updating vegetation information on land surface model performance", Journal of Geophysical Research: Atmospheres **128**, 21, 2023 039076 (2023).

Samaniego, L., S. Thober, N. Wanders, M. Pan, O. Rakovec, J. Sheffield, E. Wood, C. Prudhomme, G. Rees, H. Houghton-Carr and M. Fry, "Hydrological forecasts and projections for improved decision-making in the water sector in europe", Bulletin of the American Meteorological Society **100**, 12, 2451–2472 (2019).

Schewe, J., J. Heinke, D. Gerten, I. Haddeland, N. Arnell, D. Clark, R. Dankers, S. Eisner, B. Fekete, F. Colón-González and S. Gosling, "Multimodel assessment of water scarcity under climate change", Proceedings of the National Academy of Sciences **111**, 9, 3245–3250 (2014).

Schlesinger, W. and S. Jasechko, "Transpiration in the global water cycle", Agricultural and Forest Meteorology **189**, 115–117 (2014).

Secci, D., M. G. Tanda, M. D'Oria and V. Todaro, "Artificial intelligence models to evaluate the impact of climate change on groundwater resources", Journal of Hydrology **627**, 130359 (2023).

Sha, Z., Y. Bai, R. Li, H. Lan, X. Zhang, J. Li, X. Liu, S. Chang and Y. Xie, "The global carbon sink potential of terrestrial vegetation can be increased substantially by optimal land management", Communications Earth & Environment **3**, 1, 8 (2022).

Shen, C., "A transdisciplinary review of deep learning research and its relevance for water resources scientists", Water Resources Research **54**, 11, 8558–8593 (2018).

Shi, X., Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong and W.-c. Woo, "Convolutional lstm network: A machine learning approach for precipitation nowcasting", Advances in neural information processing systems **28** (2015).

Siebert, S., J. Burke, J.-M. Faures, K. Frenken, J. Hoogeveen, P. Döll and F. T. Portmann, "Groundwater use for irrigation–a global inventory", Hydrology and Earth System Sciences **14**, 10, 1863–1880 (2010).

Simonovic, S., "Reservoir systems analysis: closing gap between theory and practice", Journal of water resources planning and management **118**, 3, 262–280 (1992).

Singh, N. K. and N. B. Basu, "The human factor in seasonal streamflows across natural and managed watersheds of north america", Nature Sustainability **5**, 5, 397–405 (2022).

Sit, M., B. Demiray, Z. Xiang, G. Ewing, Y. Sermet and I. Demir, "A comprehensive review of deep learning applications in hydrology and water resources", Water Science and Technology **82**, 12, 2635–2670 (2020).

Song, Y., W. Tsai, J. Gluck, A. Rhoades, C. Zarzycki, R. McCrary, K. Lawson and C. Shen, "Lstm-based data integration to improve snow water equivalent prediction and diagnose error sources", Journal of Hydrometeorology (2023).

Sorkhabi, O., B. Shadmanfar and E. Kiani, "Monitoring of dam reservoir storage with multiple satellite sensors and artificial intelligence", Results in Engineering **16**, 100542 (2022).

Steyaert, J., L. Condon, W. Turner, S. and N. Voisin, "Resopsus, a dataset of historical reservoir operations in the contiguous United States", Scientific Data **9**, 1, 1–8 (2022).

Sturm, M., C. Racine and K. Tape, "Increasing shrub abundance in the arctic", Nature **411**, 6837, 546–547 (2001).

Suen, J. and J. Eheart, "Reservoir management to balance ecosystem and human needs: Incorporating the paradigm of the ecological flow regime", Water resources research **42**, 3 (2006).

Sun, L., L. Yang, L. Chen, F. Zhao and S. Li, "Hydraulic redistribution and its contribution to water retention during short-term drought in the summer rainy season in a humid area", Journal of hydrology **566**, 377–385 (2018).

Sun, R., B. Pan and Q. Duan, "A surrogate modeling method for distributed land surface hydrological models based on deep learning", Journal of Hydrology **624**, 129944 (2023).

Tang, M., Y. Liu and L. Durlofsky, "Deep-learning-based surrogate flow modeling and geological parameterization for data assimilation in 3d subsurface flow", Computer Methods in Applied Mechanics and Engineering **376**, 113636 (2021).

Tang, Q., E. Vivoni, F. Muñoz-Arriola and D. Lettenmaier, "Predictability of evapotranspiration patterns using remotely sensed vegetation dynamics during the north american monsoon", Journal of Hydrometeorology **13**, 1, 103–121 (2012).

Tesemma, Z., Y. Wei, M. Peel and A. Western, "The effect of year-to-year variability of leaf area index on variable infiltration capacity model performance and simulation of runoff", Advances in Water Resources **83**, 310–322 (2015).

Texas Water Development Board, "Volumetric and sedimentation survey of Belton Lake", Texas Water Development Board (2015).

Texas Water Development Board, University of Texas at Arlington and National Integrated Drought Information System, "Forecast-informed reservoir operation (FIRO) and water resources management in Texas", (2020).

Thorne, J., H. Choe, R. Boynton, J. Bjorkman, W. Albright, K. Nydick and M. Schwartz, "The impact of climate change uncertainty on california's vegetation and adaptation management", Ecosphere **8**, 12, 02021 (2017).

Tonkin, J., D. Merritt, J. Olden, L. Reynolds and D. Lytle, "Flow regime alteration degrades ecological networks in riparian ecosystems", Nature ecology & evolution **2**, 1, 86–93 (2018).

Tran, H., E. Leonarduzzi, L. Fuente, R. Hull, V. Bansal, C. Chennault, P. Gentine, P. Melchior, L. Condon and R. Maxwell, "Development of a deep learning emulator for a distributed groundwater–surface water model: ParFlow-ML", Water **13**, 23, 3393 (2021).

Tran, H., J. Zhang, M. O'Neill, A. Ryken, L. Condon and R. Maxwell, "A hydrological simulation dataset of the Upper Colorado River Basin from 1983 to 2019", Scientific Data **9**, 1, 1–17 (2022).

Tran, V. N., V. Y. Ivanov, D. Xu and J. Kim, "Closing in on hydrologic predictive accuracy: Combining the strengths of high-fidelity and physics-agnostic models", Geophysical Research Letters **50**, 17, e2023GL104464 (2023).

Troin, M., R. Arsenault, A. W. Wood, F. Brissette and J.-L. Martel, "Generating ensemble streamflow forecasts: A review of methods and approaches over the past 40 years", (2021).

Tsai, W.-P., D. Feng, M. Pan, H. Beck, K. Lawson, Y. Yang, J. Liu and C. Shen, "From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling", Nature communications **12**, 1, 5988 (2021).

Turner, S., J. Steyaert, L. Condon and N. Voisin, "Water storage and release policies for all large reservoirs of conterminous United States", Journal of Hydrology **603**, 126843 (2021).

Turner, S., W. Xu and N. Voisin, "Inferred inflow forecast horizons guiding reservoir release decisions across the United States", Hydrology and Earth System Sciences **24**, 3, 1275–1291 (2020a).

Turner, S. W., K. Doering and N. Voisin, "Data-driven reservoir simulation in a large-scale hydrological and water resource model", Water Resources Research **56**, 10, e2020WR027902 (2020b).

Van Den Hoek, J., A. Getirana, H. C. Jung, M. A. Okeowo and H. Lee, "Monitoring reservoir drought dynamics with landsat and radar/lidar altimetry time series in persistently cloudy eastern brazil", Remote Sensing **11**, 7, 827 (2019).

Van Dijk, A., "Model-data fusion: Using observations to understand and reduce uncertainty in hydrological models", in "19th International Congress on Modelling and Simulation", (2011), 12-16 December 2011.

Vano, J., T. Das and D. Lettenmaier, "Hydrologic sensitivities of Colorado River runoff to changes in precipitation and temperature", Journal of Hydrometeorology **13**, 3, 932–949 (2012).

Verger, A., F. Baret and M. Weiss, "GEOV2/VGT: Near real time estimation of global biophysical variables from VEGETATION-P data", in "MultiTemp 2013: 7th International Workshop on the Analysis of Multi-temporal Remote Sensing Images", p. 1–4 (IEEE, 2013).

Verger, A., F. Baret, M. Weiss, B. Smets, F. Camacho and R. Lacaze, "GEOV2/VGT: Near real time estimation of LAI, FAPAR and cover fraction variables from VEGETATION data within Copernicus Global Land service", in "Fourth International Symposium on Recent Advances in Quantitative Remote Sensing", (2014).

Vertessy, R., F. Watson and K. Sharon, "Factors determining relations between stand age and catchment water balance in mountain ash forests", Forest Ecology and Management **143**, 1-3, 13–26 (2001).

Wang, F., L. Wang, H. Zhou, O. Saavedra Valeriano, T. Koike and W. Li, "Ensemble hydrological prediction-based real-time optimization of a multiobjective reservoir during flood season in a semiarid basin with global numerical weather predictions", Water Resources Research **48**, 7 (2012).

Wang, H., X. Huo, Q. Duan, R. Liu and S. Luo, "Uncertainty quantification for the noah-mp land surface model: A case study in a grassland and sandy soil region", Journal of Geophysical Research: Atmospheres **128**, 20, 2023 038556 (2023).

Wang, L., S. Good and K. Caylor, "Global synthesis of vegetation control on evapotranspiration partitioning", Geophysical Research Letters **41**, 19, 6753–6757 (2014).

Ward, S., D. Scott Borden, A. Kabo-Bah, A. N. Fatawu and X. F. Mwinkom, "Water resources data, models and decisions: international expert opinion on knowledge management for an uncertain but resilient future", Journal of Hydroinformatics **21**, 1, 32–44 (2019).

Wattenbach, M., D. Franz, W. Liang, M. Schmidt, F. Seitz and A. Güntner, "Integration of modis lai products into the hydrological model wghm indicate the sensitivity of total water storage simulations to vegetation cover dynamics", in "EGU General Assembly Conference Abstracts", p. 10116 (2012-04).

Webster, C. and T. Jonas, "Influence of canopy shading and snow coverage on effective albedo in a snow-dominated evergreen needleleaf forest", Remote Sensing of Environment **214**, 48–58 (2018).

Wei, C. and N. Hsu, "Derived operating rules for a reservoir operation system: Comparison of decision trees, neural decision trees and fuzzy decision trees", Water resources research **44**, 2 (2008).

Wei, H., Y. Xia, K. Mitchell and M. Ek, *Improvement of Noah land surface model for the warm season processes: Assessment of water and energy flux simulation* (Hydrol. Processes, 2012).

Wei, P., Z. Lu and J. Song, "Variable importance analysis: A comprehensive review", Reliability Engineering & System Safety **142**, 399–432 (2015).

Woodhouse, C., S. Gray and D. Meko, "Upyeard streamflow reconstructions for the Upper Colorado River Basin", Water Resources Research **42**, 5 (2006).

Wurbs, R., "Storage and regulation of river flows by dams and reservoirs", Texas Water Journal **12**, 1, 10–39 (2021).

Wurbs, R., P. Carriere and W. Johnson, "Management strategies for increasing reservoir yields", Water international **15**, 3, 168–175 (1990).

Wurbs, R., R. Muttiah and F. Felden, "Incorporation of climate change in water availability modeling", Journal of Hydrologic Engineering **10**, 5, 375–385 (2005).

Xia, Y., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E. Wood, L. Luo, C. Alonge, H. Wei, J. Meng *et al.*, "Continental-scale water and energy flux analysis and validation for the north american land data assimilation system project phase 2 (NLDAS-2): 1. intercomparison and application of model products", Journal of Geophysical Research: Atmospheres **117**, D3 (2012).

Xiao, M., B. Udall and D. Lettenmaier, "On the causes of declining Colorado River streamflows", Water Resources Research **54**, 9, 6739–6756 (2018).

Xu, T. and F. Liang, "Machine learning for hydrologic sciences: An introductory overview", Wiley Interdisciplinary Reviews: Water **8**, 5, e1533 (2021).

Xu, T. and A. J. Valocchi, "A bayesian approach to improved calibration and prediction of groundwater models with structural error", Water Resources Research **51**, 11, 9290–9311 (2015).

Xu, T., A. J. Valocchi, M. Ye, F. Liang and Y.-F. Lin, "Bayesian calibration of groundwater models with input data uncertainty", Water Resources Research **53**, 4, 3224–3245 (2017).

Yan, K., T. Park, G. Yan, C. Chen, B. Yang, Z. Liu, R. Nemani, Y. Knyazikhin and R. Myneni, "Evaluation of modis lai/fpar product collection 6. part 1: Consistency and improvements", Remote Sensing **8**, 5, 359 (2016).

Yang, S., D. Yang and J. Chen, "A machine-learning based surrogate model for distributed hydrological simulation", in "AGU Fall Meeting Abstracts", vol. 2019, pp. H53K–1916 (2019a).

Yang, T., A. Asanjan, E. Welles, X. Gao, S. Sorooshian and X. Liu, "Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information", Water Resources Research **53**, 4, 2786–2812 (2017).

Yang, T., X. Gao, S. Sorooshian and X. Li, "Simulating california reservoir operation using the classification and regression-tree algorithm combined with a shuffled cross-validation scheme", Water Resources Research **52**, 3, 1626–1651 (2016).

Yang, T., L. Zhang, T. Kim, Y. Hong, D. Zhang and Q. Peng, "A large-scale comparison of artificial intelligence and data mining (ai&dm) techniques in simulating reservoir releases over the Upper Colorado region", Journal of Hydrology **602**, 126723 (2021).

Yang, Y., M. Pan, H. Beck, C. Fisher, R. Beighley, S. Kao, Y. Hong and E. Wood, "In quest of calibration density and consistency in hydrologic modeling: Distributed parameter calibration against streamflow characteristics", Water Resources Research **55**, 9, 7784–7803 (2019b).

Yassin, F., S. Razavi, M. Elshamy, B. Davison, G. Sapriza-Azuri and H. Wheater, "Representation and improved parameterization of reservoir operation in hydrological and land-surface models", Hydrology and Earth System Sciences **23**, 9, 3735–3764 (2019).

Yates, D., J. Sieber, D. Purkey and A. Huber-Lee, "WEAP21—A demand-, priority-, and preference-driven water planning model: part 1: model characteristics", Water international **30**, 4, 487–500 (2005).

Yeo, I., J. Guldmann and S. Gordon, "A hierarchical optimization approach to watershed land use planning", Water resources research **43**, 11 (2007).

Yin, J., X. Zhan, Y. Zheng, C. Hain, M. Ek, J. Wen, L. Fang and J. Liu, "Improving noah land surface model performance using near real time surface albedo and green vegetation fraction", Agricultural and forest meteorology **218**, 171–183 (2016).

Yu, K. and P. D'Odorico, "Climate, vegetation, and soil controls on hydraulic redistribution in shallow tree roots", Advances in Water Resources **66**, 70–80 (2014).

Yuan, W., Y. Zheng, S. Piao, P. Ciais, D. Lombardozzi, Y. Wang, Y. Ryu, G. Chen, W. Dong, Z. Hu and A. Jain, "Increased atmospheric vapor pressure deficit reduces global vegetation growth", Science advances **5**, 8, 1396 (2019).

Zahura, F. and J. Goodall, "Predicting combined tidal and pluvial flood inundation using a machine learning surrogate model", Journal of Hydrology: Regional Studies **41**, 101087 (2022).

Zahura, F., J. Goodall, J. Sadler, Y. Shen, M. Morsy and M. Behl, "Training machine learning surrogate models from a high-fidelity physics-based model: Application for real-time street-scale flood prediction in an urban coastal community", Water Resources Research **56**, 10, 2019 027038 (2020).

Zarei, M., O. Bozorg-Haddad, S. Baghban, M. Delpasand, E. Goharian and H. A. Loáiciga, "Machine-learning algorithms for forecast-informed reservoir operation (FIRO) to reduce flood damages", Scientific reports **11**, 1, 24295 (2021).

Zeng, R. and W. Ren, "The spatiotemporal trajectory of us agricultural irrigation withdrawal during 1981–2015", Environmental Research Letters **17**, 10, 104027 (2022).

Zhang, B., Y. Xia, B. Long, M. Hobbins, X. Zhao, C. Hain, Y. Li and M. Anderson, "Evaluation and comparison of multiple evapotranspiration data models over the contiguous United States: Implications for the next phase of NLDAS (NLDAS-Testbed) development", Agricultural and Forest Meteorology **280**, 107810 (2020a).

Zhang, D., J. Lin, Q. Peng, D. Wang, T. Yang, S. Sorooshian, X. Liu and J. Zhuang, "Modeling and simulating of reservoir operation using the artificial neural network, support vector regression, deep learning algorithm", Journal of Hydrology **565**, 720–736 (2018).

Zhang, Y., P. Gentine, X. Luo, X. Lian, Y. Liu, S. Zhou, A. Michalak, W. Sun, J. Fisher, S. Piao and T. Keenan, "Increasing sensitivity of dryland vegetation greenness to precipitation due to rising atmospheric CO_2", Nature communications **13**, 1, 1–9 (2022).

Zhang, Y., N. Parazoo, A. Williams, S. Zhou and P. Gentine, "Large and projected strengthening moisture limitation on end-of-season photosynthesis", Proceedings of the National Academy of Sciences **117**, 17, 9216–9222 (2020b).

Zhang, Y., K. Wang, J. Wang, C. Liu and Z. Shangguan, "Changes in soil water holding capacity and water availability following vegetation restoration on the chinese loess plateau", Scientific Reports **11**, 1, 9692 (2021).

Zhao, Q. and X. Cai, "Deriving representative reservoir operation rules using a hidden markov-decision tree model", Advances in Water Resources **146**, 103753 (2020).

Zhao, T., X. Cai and D. Yang, "Effect of streamflow forecast uncertainty on real-time reservoir operation", Advances in water resources **34**, 4, 495–504 (2011).

Zhao, T., J. Zhao, J. Lund and D. Yang, "Optimal hedging rules for reservoir flood operation from forecast uncertainties", Journal of Water Resources Planning and Management **140**, 12, 04014041 (2014).

Zhao, W., X. Yu, C. Xu, S. Li, G. Wu and W. Yuan, "Dynamic traceability effects of soil moisture on the precipitation–vegetation association in drylands", Journal of Hydrology **615**, 128645 (2022).

Zhao, Y., S. Liu and H. Shi, "Impacts of dams and reservoirs on local climate change: a global perspective", Environmental Research Letters **16**, 10, 104043 (2021).

Zhu, J., Y. Yang, H. Holmquist and N. Leber, "Projected reservoir rating curves based on sedimentation surveys and its application in water planning in Texas", in "World Environmental and Water Resources Congress 2018", pp. 349–362 (American Society of Civil Engineers Reston, VA, 2018a).

Zhu, X., F. Cai, J. Tian and T. Williams, "Spatiotemporal fusion of multisource remote sensing data: Literature survey, taxonomy, principles, applications, and future directions", Remote Sensing **10**, 4, 527 (2018b).

Zhu, Z., J. Bi, Y. Pan, S. Ganguly, A. Anav, L. Xu, A. Samanta, S. Piao, R. Nemani and R. Myneni, "Global data sets of vegetation leaf area index (LAI) 3g and fraction of photosynthetically active radiation (FPAR) 3g derived from global inventory modeling and mapping studies (GIMMS) normalized difference vegetation index (NDVI3g) for the period 1981 to 2011", Remote sensing **5**, 2, 927–948 (2013).

APPENDIX A

DATA SOURCES

For the Chapter 2, the meteorological forcing (precipitation, potential evapotranspiration, and air temperature) is available at https://ldas.gsfc.nasa.gov/nldas/v2/forcing. Snow depth data is retrieved from Daily 4 km Gridded SWE and Snow Depth from Assimilated In Situ and Modeled Data over the Conterminous US, Version 1 (NSIDC-0719) (https://nsidc.org/data/nsidc-0719/versions/1). The data set of reservoir operations is available online (https://www.hydroshare.org/resource/79c262b627fc4ce293379b5e95457146/), or directly from the United States Bureau of Reclamation (https://water.usbr.gov/api/web/app.php/api/) and the United States Army Corps of Engineers (collected via Duke University; https://nicholasinstitute.duke.edu/reservoir-data/, Patterson and Doyle, 2018).

For the Chapter 3, the data set utilized in this dissertation can be assessed online (https://www.hydroshare.org/resource/f32010b5edc34e519399e2417db2cce6/).

For the Chapter 4, the data set utilized in this dissertation can be assessed online (https://www.hydroshare.org/resource/e5ee0fec0fd444558c6104997a65a4d6/).