From Sources to Sinks: Advancing Surface Water Management Through Satellite Remote Sensing

Shahzaib Khan

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2026

Reading Committee:

Faisal Hossain, Chair

Erkan Istanbulluoglu

Nicoleta Cristea

Program Authorized to Offer Degree:

Civil and Environmental Engineering

University of Washington

Abstract

From Sources to Sinks: Advancing Surface Water Management Through Satellite Remote Sensing

Shahzaib Khan

Chair of the Supervisory Committee:
Professor Faisal Hossain
Civil and Environmental Engineering

Surface water is essential for sustaining life, supporting economic activity, and ensuring food security. Historically, since the dawn of remote sensing from space, we have been able to determine where water is located and its spatial extent (Area - L^2). However, for deeper process-based insights and to inform decision-making, knowledge of spatial extent is not sufficient. We require knowledge of volume (L^3) and fluxes (L^3/T) to understand the fate and transport of surface water. This challenge motivates this study, which seeks to answer the

overarching question: ***How can satellite remote sensing be utilized to enhance surface water management?***

To address this question, the research develops a set of integrated frameworks based on a *Source-to-Sink* (S2S) perspective, monitoring both sources, where surface water is stored (e.g., lakes and wetlands), and sinks, where it is consumed (e.g., irrigated agriculture). The first component integrates satellite remote sensing and citizen science to develop an innovative, cost-effective methodology for monitoring and managing the volume and flux of surface water. A time-averaged uncertainty metric revealed higher errors in mountainous regions due to shadowing and pixel misclassification. This approach establishes a pre-SWOT (Surface Water and Ocean Topography satellite mission) benchmark for satellite-based monitoring and demonstrates how quantified uncertainty can inform water management decisions, such as designing urban water supply systems for minimum lake volumes and flood infrastructure for maximum levels.

Building on this work, the monitoring strategy for surface water is advanced by developing a gauge network optimization framework that integrates hydro-physical factors such as temperature, precipitation, and terrain roughness, along with citizen science elevation data. The framework introduces two novel methods to model lake area-volume relationships, providing benchmark volume estimates for regional and global assessments. A probabilistic mapping approach identifies optimal gauge placements to maximize representativeness and return on investment. Furthermore, the framework supports the integration of satellite-derived virtual stations and facilitates future expansion of gauge networks in coordination with missions such as SWOT. The analysis also showed that strategically deploying around 30 gauges across one-fourth of Bangladesh (Northeastern Bangladesh, characterized by flat topography) could significantly enhance surface water monitoring through more frequent and representative observations in the region.

Extending the *S2S* framework to the domain of water consumption, the next component of this research focuses on irrigation, which is the largest consumer of global water resources. A decision-support framework for gravity-fed canal systems was developed to assist water managers (water providers) in allocating water based on crop-specific demand. By integrating Earth observation data with global numerical weather models, the framework quantifies the net water requirements for each canal within the network. The results demonstrate that satellite-informed allocation can significantly enhance irrigation efficiency, reducing water supply requirements by approximately 26% of the supply.

Finally, an open-source Python package that unites both water providers and consumers within a single digital umbrella. Building on earlier efforts that focused primarily on providers, this package integrates satellite, in-situ, and local weather station data thereby extending beyond reliance on global models. Its key contribution is its design philosophy that emphasizes transparency, scalability, and accessibility through the use of open-source tools, publicly available datasets, detailed documentation, and flexible user configurations, ultimately lowering barriers to adoption across diverse agricultural and technological contexts.

Collectively, this research highlights the potential of coupling satellite remote sensing with citizen science and cloud-based platforms for quantifying surface water storage (L^3) and enhancing surface water management. Developed in consultation with stakeholders, the frameworks operationalize the *S2S* approach, bridging scientific innovation and real-world implementation. The results offer a scalable pathway toward more equitable, efficient, and sustainable surface water management in a changing world.

ACKNOWLEDGEMENTS

First and foremost, I offer my heartfelt gratitude to Almighty God for His boundless compassion, guidance, and care throughout my life. His blessings have led me to this point and surrounded me with supportive friends, colleagues, and mentors who have contributed to my growth as a researcher and as a person.

I am deeply grateful to Professor Faisal Hossain, whose mentorship has been transformative. His guidance, patience, and unwavering encouragement have shaped this research and my growth as a scientist. Working with him has been an extraordinary privilege as he embodies a rare balance of encouragement and rigor - never hesitating to point out mistakes, challenge assumptions, or push for deeper clarity, yet always with the pure-hearted intention of helping me grow and build a successful future.

I am also thankful to my committee members, Erkan Istanbulluoglu and Nicoleta Cristea, for their thoughtful feedback, encouragement, and support, which have enriched this work.

I extend my appreciation to the LOCSS team from the University of North Carolina and Tennessee Tech University for their constant support and collaboration. I am equally grateful to my fellow SASWE lab mates at the University of Washington, whose camaraderie, encouragement, and shared curiosity made this journey both stimulating and enjoyable, as well as to all my office mates in the Fishbowl and Treehouse.

Finally, I cannot express enough gratitude to my family and friends. To my parents, both dedicated educators, and my siblings, your unconditional love, patience, and faith in me have been my strongest support. From my parents, I learned the value of working selflessly to benefit humanity, a lesson that has guided my academic journey and personal growth. Your encouragement has carried me through every challenge, and your unwavering belief in my potential has been a constant source of inspiration.

DEDICATION

This thesis is dedicated to my parents, Mr. Shakeel Khan Pathan and Mrs. Tabassum, and my siblings, Dr. Aayaisha Khan and Dr. Alveena Khan.

Published Material

At the time of writing, Chapter 2 has been published in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* Chapter 3 has been published in the *Environmental Modeling and Software* journal. Chapter 4 has been submitted to *Hydrology and Earth System Sciences (HESS)*. Chapter 5 has been submitted to *Digital Water*. These articles are:

Chapter 2:

S. Khan *et al*., "Understanding Volume Estimation Uncertainty of Lakes and Wetlands Using Satellites and Citizen Science," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 2386-2401, 2023, doi: 10.1109/JSTARS.2023.3250354.

Chapter 3:

Khan, S., Hossain, F., Pavelsky, T., Gomez, A., Ghafoor, S., Lane, M., Parkins, G., Minocha, S., Bhuyan, M. A., Al Fayyaz, T. A., Haque, M. N., Sarker, P. K., & Borua, P. P. (2024). A network design approach for citizen science-satellite monitoring of surface water volume changes in Bangladesh. *Environmental Modelling & Software*, *172*, 105919. https://doi.org/10.1016/j.envsoft.2023.105919

Chapter 4:

Khan, S., Hossain, F., Islam, K., and Ahamed, M. (2025). Satellite Data Rendered Irrigation using Penman-Monteith and SEBAL (sDRIPS) for Surface Water Irrigation Optimization, *Hydrology and Earth System Sciences (HESS),* preprint - https://doi.org/10.5194/egusphere-2025-4574

Chapter 5:

Khan, S., Hossain, F. (2025). (sDRIPS): A Cloud-Based, Open-Source Python Package for Satellite-Informed Surface Water Irrigation Optimization. *Digital Water*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1. Introduction

1.1. Background Information

1.1.1. Human Dependence on Water and Food

Water and food are the most fundamental requirements for sustaining human life. Every facet of human development - health, economic growth, energy production, environmental and political stability depends on the availability of freshwater. Although water covers approximately 71% of Earth's surface, only about 2.5% of this is freshwater, and less than 1% is readily accessible in lakes, rivers, and wetlands (Gleick, 2014). This small yet vital fraction sustains not only humans but also ecosystems that maintain biodiversity and regulate hydrological and climatic processes.

The spatial and temporal distribution of freshwater is highly uneven, creating persistent regional disparities in access and availability. Throughout human history, the presence of freshwater has determined where civilizations emerge, thrive, or decline. Today, the same dependency remains unchanged: freshwater underpins agriculture, industry, and ecosystem health. Agriculture alone consumes roughly 85% of global water withdrawals (D'Odorico et al., 2020), linking water scarcity directly to food insecurity and economic vulnerability. The stability of freshwater systems, therefore, is not merely an environmental concern, it is a defining factor for societal resilience and sustainable development.

1.1.2. Mounting Stress on Global Water Resources

Despite its dependence, freshwater systems are under increasing stress. Over the past few decades, both natural and human-induced factors have altered hydrological regimes at unprecedented rates (Kuang et al., 2024; J. Liu et al., 2019; Padrón et al., 2020). Downing et al. (2006) estimated that the total surface area of natural and artificial lakes exceeds 4.6 million km², while Verpoorter et al. (2014) suggested that there are approximately 117 million lakes globally. However, numerous studies have documented a persistent decline in surface water extent across continents in recent decades (Courouble et al., 2021; Wine & Laronne, 2020), signaling widespread contraction of global freshwater bodies.

The human implications of this contraction are profound. Mekonnen & Hoekstra, (2016) estimated that two-thirds of the global population (around 4.0 billion people) experience severe water scarcity for at least one month each year, while half a billion face such scarcity throughout the year. Without significant improvements in water-use efficiency, this stress is expected to intensify further. Projections from the World Resources Institute (https://www.wri.org/, Figure 1.1) show that 51 of the 164 countries and territories are expected to suffer from high to extremely high water stress by 2050. In addition to the entire Arabian Peninsula, Iran and India, most North African countries such as Algeria, Egypt and are expected to consume at least 80% (shown in greyed brackets) of their available water resources by 2050.

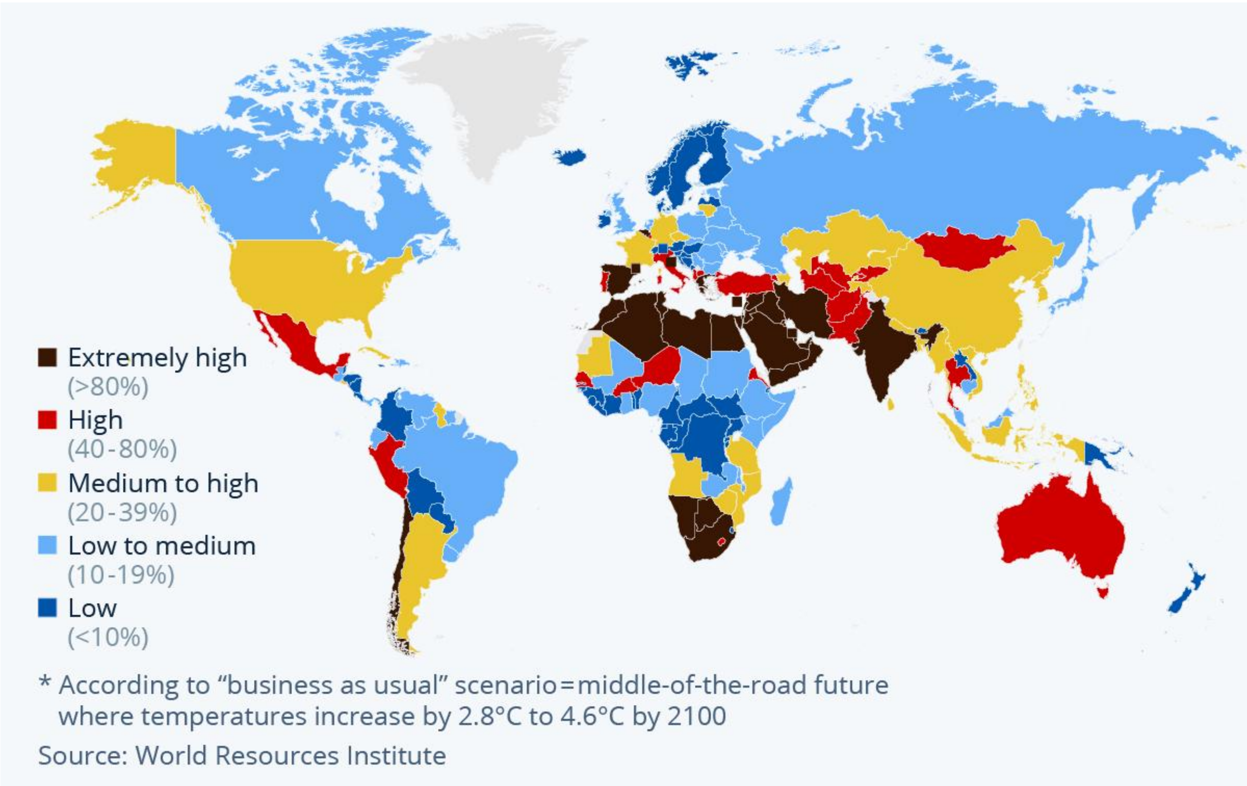

Figure 1.1: Representation of global water stress level (the proportion of water withdrawn for use in industry, agriculture or private households in relation to available water). Source - *World Resources Institute (https://www.wri.org/applications/aqueduct/country-rankings/?indicator=bws)*.

Compounding this pressure is the mounting demand for food. Feeding a global population projected to exceed nine billion by 2050 will require substantial increases in food production, placing even greater strain on freshwater systems (Boretti & Rosa, 2019). This coupling of water and food security underscores the urgent need for sustainable and adaptive management of freshwater resources. Achieving such a goal first requires a critical understanding of how water resources are currently monitored, what their limitations are, and how emerging technologies might bridge these gaps to ensure long-term water security.

1.1.3. Current Status of Monitoring

Effective management of surface water hinges on robust monitoring networks capable of capturing hydrological dynamics across spatial and temporal scales. Historically, in-situ gauge networks have served as the backbone of any hydrologic observation systems, providing long-term records of river discharge, lake levels, and precipitation. However, these networks have declined globally over recent decades due to financial constraints, limited institutional support, and inadequate maintenance (Burn, 1997; Mishra & Coulibaly, 2009; Pilon et al., 1996; Soares & Calijuri, 2021). In addition, conflicts, natural disasters, and pandemics have also disrupted monitoring programs, resulting in spatially biased networks with inadequate coverage, particularly in developing regions across the globe.

Satellite remote sensing has greatly expanded observation capacity, yet significant limitations persist. Until the launch of the Surface Water and Ocean Topography (SWOT) satellite in late 2022, no global mission could simultaneously observe water surface area and elevation. Previous satellites could only provide either spatial extent (L^2) using optical and radar imagery, or elevation (L) using altimetry, each with limited temporal frequency and spatial coverage. The SWOT mission now enables the simultaneous estimation of water surface area and elevation, marking a critical advancement in global hydrologic observation. Nevertheless, deriving accurate volumetric storage (L^3) remains a major challenge due to the lack of bathymetric data. Moreover, all remote sensing approaches whether optical, radar, altimetric, or interferometric, introduce uncertainties and face challenges in observing small or dynamic water bodies in complex environment. Integration of satellite-derived estimates with in situ measurements is therefore essential to reduce uncertainty and advance hydrologic monitoring.

Parallel to surface water monitoring, agricultural water management presents another layer of complexity. Efficient irrigation scheduling and allocation require timely and spatially explicit data on soil moisture, evapotranspiration, and crop water demand. Yet, traditional irrigation practices often rely on fixed schedules decisions on *when, how much and where* to irrigate based on experience or intuition rather than data-driven insight. While advisory systems have been developed to improve irrigation management, they remain limited by high costs, dense sensor requirements, and restricted scalability (Chen et al., 2021; Forcén-Muñoz et al., 2022; Sishodia et al., 2020).

Satellite-based products such as Moderate Resolution Imaging Spectroradiometer (MODIS), ECOSTRESS, and Landsat have been widely used to generate irrigation advisories. However, these systems are constrained by the trade-off between spatial and temporal resolution. High-temporal-resolution data (daily) typically come at coarser spatial scales (kilometers), whereas higher spatial resolution data (tens of meters) are available only at multi-day to biweekly intervals (Chen et al., 2021; Tasumi, 2019). Although ECOSTRESS provides evapotranspiration estimates at 70 m resolution, its operational consistency has been affected by intermittent data availability and thermal sensor failure (Hulley et al., 2022). These trade-offs make it challenging to monitor crop water use continuously at the field scale, and allocate water based on the actual requirements.

Remote sensing advances have nonetheless transformed our ability to locate and characterize water bodies. Global datasets now map the areal extent (L^2) of millions of surface water bodies with unprecedented precision (Hu et al., 2017; Lehner & Döll, 2004; Messager et al., 2016; Pekel et al., 2016; Verpoorter et al., 2014). However, the global distribution of surface water volume (L^3) remains largely unknown. For effective water management, spatial extent alone is insufficient; understanding the volume (L^3) of stored water and the fluxes (L^3/T) that govern its movement is critical for assessing availability, predicting scarcity, and guiding sustainable allocation. Monitoring systems must therefore evolve beyond mapping to quantifying storage and consumption, particularly in agriculture, the dominant consumer of global freshwater resources.

1.1.4. Sources, Sinks, and the Need for Integrated Management

As the preceding discussion highlights, the global water crisis is not solely a challenge of scarcity but also one of limited information on its availability (L^3), fragmented monitoring, and inefficient management practices. Addressing this crisis requires solutions that are both scientifically robust and operationally sustainable. Effective water management, therefore, demands a *Source-to-Sink (S2S)* approach that considers the entire hydrologic continuum from water storage to water use.

In this *S2S* framework, *sources* refer to the reservoirs of freshwater systems, such as lakes, rivers, and wetlands, that store and supply water. *Sinks,* on the other hand, represent the sectors and processes that consume water, particularly irrigated agriculture, which accounts for the vast majority of withdrawals globally. Managing one without the other provides only a partial understanding of water sustainability. The dynamics of sources, such as how their storage volumes fluctuate spatially and temporally, must be critically evaluated alongside the uncertainties and challenges involved in their observation. Similarly, understanding sinks requires monitoring how water is allocated and consumed across agricultural systems, where inefficiencies and data limitations often obscure the true magnitude of water use.

A comprehensive management framework must therefore integrate both the dimensions, recognizing that the sustainability of one inherently depends on the balance and feedbacks with the other (*source-sink*). Transitioning towards this integration requires quantifying how much water can be withdrawn from sources and how much can be consumed within sinks without depleting or degrading the system. Despite remarkable advances in Earth observation, hydrological modeling, and sensor technology, translating scientific knowledge into actionable management strategies remains limited. Bridging this divide demands both scientific integration through data fusion, uncertainty reduction, and coupled monitoring systems, and science-society interaction where research outputs are made accessible, interpretable, and actionable for decision-makers.

Citizen science (collecting data through volunteer citizens) represents a promising avenue for strengthening this interface between research and practice. By involving local communities in data collection and validation, citizen science can fill observational gaps, and foster stewardship of water resources. Harnessing the combined strengths of satellite observations, citizen science, and cloud-based data infrastructures provides an integrated, scalable, and transparent foundation for sustainable water management. This *S2S* oriented framework empowers researchers to better understand hydrologic processes while enabling water managers to make data-driven, equitable, and adaptive decisions that balance the needs of both sources and sinks.

1.1.5. Research Objectives

Building upon this vision, this dissertation seeks to advance the understanding, monitoring, and management of surface water resources through the synergistic use of satellite remote sensing, in-situ observations, citizen science data, and cloud computing. It is guided by the central research question - ***How can the integration of satellite remote sensing and ground-based observations enhance the monitoring and management of surface water resources?*** By improving water resource management, this work aims to contribute to Sustainable Development Goals (SDG) 13 (Climate Action) by enabling adaptive, data-driven strategies to mitigate climate-induced water stress. The specific research objectives are as follows:

1. Quantify uncertainties in estimating the volume of surface water stored in lakes, and wetlands using satellite remote sensing and citizen science gauge data, and evaluate how these uncertainties vary with geography, season, and environmental conditions.
2. Optimize surface water monitoring networks by integrating existing gauges with current and future satellite missions to improve volume estimation and spatial representativeness.
3. Optimize multi-branched canal irrigation systems through Earth observation data, aligning water delivery with spatiotemporal crop water requirements to improve allocation efficiency.
4. Develop a standalone open-source Python package that supports both water providers and consumers in data-driven water allocation, integrating satellite products, in-situ measurements, and local meteorological data into a unified decision-support tool.

In the long term, this dissertation aims to form a cohesive *S2S* framework that bridges fundamental understanding and real-world applications. By integrating science, technology, and citizen science, the work aspires to promote a water-secure and sustainable future. Stakeholder experience and iterative feedback are embedded in the research, ensuring that solutions are locally relevant and scalable to other regions facing similar water management challenges.

1.2. Approach

To develop the dual focused monitoring (*S2S*) frameworks proposed in this dissertation, the research is distributed into four core chapters. The first two chapters concentrate on monitoring and quantifying the sources of freshwater, such as lakes, wetlands, and reservoirs, while the latter two focus on monitoring and managing the sinks, with particular emphasis on irrigated agriculture. This structure ensures a coherent transition from understanding water availability to optimizing water use, reflecting the interconnected nature of supply and demand of our hydrologic systems.

Chapter 2 focuses on quantifying the uncertainties associated with estimating surface water volume using multi-satellite observations and gauge data collected through citizen science. The chapter systematically evaluates how these uncertainties vary across geographical, seasonal, and environmental contexts. Moreover, it establishes a benchmark between the state-of-the-art techniques available at the time of the study (2021-2022) and future satellite missions such as SWOT (as per the time of study, 2021-2022), which are expected to substantially reduce these uncertainties.

Chapter 3 builds upon this foundation by developing a scalable optimization framework for surface water monitoring gauge networks. By integrating existing gauge observations with current and forthcoming satellite missions, the framework enhances the spatial representativeness, efficiency, and resilience of hydrologic observation systems, supporting long-term monitoring.

Transitioning to the sinks, Chapter 4 introduces a decision-support framework to guide canal operators or irrigation managers in allocating water efficiently across multi-branched canal systems. Leveraging Earth observation data, evapotranspiration models, and numerical weather models, this framework enables dynamic, data-informed irrigation scheduling that minimizes inefficiencies and aligns water delivery with the actual crop demand.

Finally, Chapter 5 translates these scientific and technical advances into a standalone open-source Python package, integrating satellite data, in-situ measurements, and local weather information within a cloud-based environment. The package is designed to lower the barrier to entry for researchers, practitioners, and water managers by leveraging publicly available datasets and open-source tools. In doing so, it bridges the gap between scientific innovation and field-level implementation, serving both water providers and consumers. Ultimately, the package promotes data-driven, transparent, and scalable decision-making for sustainable water management across diverse hydrologic and agricultural contexts.

Together, these chapters establish a unified *S2S* framework that bridges hydrologic science, technology, and participatory engagement. By jointly addressing the *sources* and *sinks,* the dissertation advances a resilient, inclusive, and technology-enabled approach to water resource monitoring and management.

Chapter 2. Understanding Volume Estimation Uncertainty of Lakes and Wetlands Using Satellites and Citizen Science

2.1. Abstract

We studied variations in the volume of water stored in small lakes and wetlands using satellite remote sensing and lake water height data contributed by citizen scientists. A total of 94 water bodies across the globe were studied using satellite data in the optical and microwave wavelengths from Landsat 8, Sentinel-1, and Sentinel-2. The uncertainty in volume estimation as a function of geography and geophysical factors, such as cloud cover, precipitation, and water surface temperature, was studied. The key finding that emerged from this global study is that uncertainty is highest in regions with a distinct precipitation season, such as in the monsoon dominated South Asia or the Pacific Northwestern region of the USA. This uncertainty is further compounded when small lakes and wetlands are seasonal with alternating land use as a water body and agricultural land, such as the wetlands of Northeastern Bangladesh. On an average, 45% of studied lakes could be estimated of their volume change with a statistical significant uncertainty that is less than the expected volume in South Asia. In North America, this statistically significant uncertainty in volume estimation was found to be around 50% in lakes eastward of the 108th meridian with lowest uncertainty found in lakes along the East coast of the USA. The article provides a baseline for understanding the current state of the art in estimating volumetric change of lakes and wetlands using citizen science in anticipation of the recently launched Surface Water and Ocean Topography Mission.

Note: This chapter has been published in its current form as an article in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (Khan et al., 2023). The only differences are in the numbering of figures, tables, and references to other sections.

2.2. Introduction

Water bodies such as small lakes (i.e. those smaller than 100 km^2) and wetlands provide vital functions for ecosystems and sustain biodiversity. Globally, wetlands cover an area of 1.2 billion hectares, which is equivalent to the area of Canada (Courouble et al., 2021). Downing et al., (2006) claimed that the total surface area of natural and artificial lakes is over 4.6 million km^2 which translates to about 117 million (Verpoorter et al., 2014). These water bodies act as biological supermarkets and groundwater recharge and discharge points, and they provide both water and nutrients necessary for crop production. Wetlands and small lakes also help in flood control and support eco-tourism. According to Global Wetland Outlook (Ramsar Convention, 2021), wetlands have been rapidly declining. Approximately 35% of the world's wetlands have disappeared since 1970 (Courouble et al., 2021). While there are various physical drivers that affect the behavior of wetlands and small lakes, the most critical among them, other than perhaps direct human management, is likely changing patterns of weather, hydrology and climate (Courouble et al., 2021).

In recent years, our ability to track the extent of small lakes and wetlands has increased manifold. Lehner & Döll, (2004) developed the Global Lakes and Wetlands Database (GLWD), which provides the maps and water surface area of the lake, wetlands, reservoirs and rivers. For lakes, GLWD was superseded by the HydroLAKES database (Messager et al., 2016), which mapped about 1.4 million lakes larger than 0.1 km2 (see Figure 2.1). A study by (Sheng et al., 2016) mapped 7.7 million lakes that are larger than 0.004km2. Meanwhile, (Verpoorter et al., 2014) used satellite imagery to identify more than 117 million lakes globally. (Hu et al., 2017) studied the areal extent of the wetlands and lakes by developing a new index (Precipitation Topographic Wetness Index). However, despite this improved understanding of the location and extent of small lakes and wetlands, understanding of the physical behavior of wetlands and lakes around the world remains limited, especially in developing regions. Specifically, we understand very little about how volumetric changes of lakes and wetlands modulate over time and as a function of climate, season or geographic region. A primary reason for this gap is the paucity of in-situ lake and wetland gauges relative to the widespread presence of lakes and wetlands around the globe. Unlike for rivers, there are no major national or international repositories dedicated to storing in situ lake level data. Even in developed countries, such data is limited—for example, the US Geological Survey gauges more than ten thousand rivers but only a few hundred lakes.

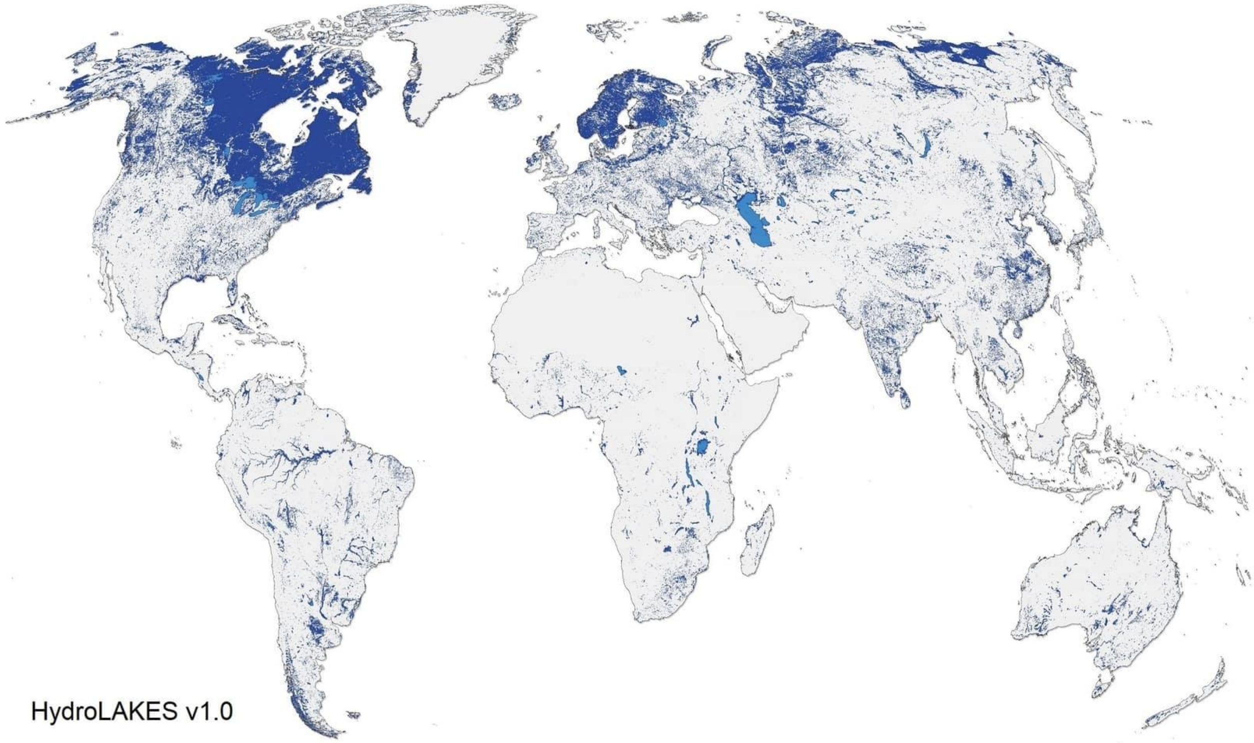

Figure 2.1: HydroLAKES Database Depiction of Global Lakes and Wetlands

Studying volumetric changes at a global scale is therefore not feasible using limited in-situ gauges given the remoteness of numerous water bodies and lack of economic or institutional resources to maintain an in-situ measurement network. Studying wetlands and small lakes using satellite remote sensing is a only cost-effective and feasible way to understand the volumetric change of water bodies on a global scale (Alsdorf & Lettenmaier, 2003; Lettenmaier et al., 2015). Most studies that aim to do so use different sensors to study surface water extent and water surface elevation, which, in combination, allow estimation of volume change. Detection of the surface water extent and elevation can be performed with sensors of different resolutions and electromagnetic wavelengths. Coarser spatial resolution sensors such as NOAA/AVHRR and Moderate Resolution Imaging Spectroradiometer (MODIS) have low spatial accuracy but high temporal resolution and coverage and are often used to study large lakes (McCullough et al., 2012). Medium spatial resolution sensors, with a resolution of around 10m - 30m and are widely used in studies of smaller lakes (Verpoorter et al., 2014). A few examples for medium resolution are the Landsat series, Sentinel 2 and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). High spatial resolution sensors such as Planet, RapidEye, IKONOS have a resolution around 1m - 5m, but they are not freely available. The type and nature of water bodies that can be studied with reasonable accuracy usually depend on the pertinent resolution and sampling frequency of sensor data that is available.

2.2.1. Research Question

Our study explores uncertainty in volume estimation, which can provide valuable information to the decision-makers or stakeholders to make more robust decisions based on estimation uncertainty. There are various factors that affect the uncertainty in volume estimation. For example, in South Asia, a key source of uncertainty is likely to be cloud cover during monsoon for the optical sensors and inundated vegetation for SAR microwave sensors. At higher latitudes or mountainous regions where lakes freeze, the area estimation may be more challenging due to the limitations of detecting inundation variations due to ice cover or due to the shadow effect of the high topography.

In this study, we have explored four different techniques that monitor inundation extent, and thus estimate volume (Table 2.1). The key research question being addressed is - ***what is the range of uncertainty associated with estimating the volume of lakes and wetlands using current sensors, and how does this uncertainty vary as a function of geography, season and average environmental conditions?*** We used data from 106 lakes and wetlands, in which water level changes were monitored by LOCSS citizen scientists. Validation of water levels collected by citizen scientists against automated water level gauges shows that they are highly accurate, with uncertainties of less than 2 cm (Little et al., 2021). Such high performance in lake level estimation can be achieved only for very large lakes using satellite altimetry. We have also used data from non-citizen programs (such as automatic gauging) when necessary to fill in gaps in our lake water height database.

Table 2.1: Summary of Sensors And Techniques Used For Lake Area Estimation

Sensor	Revisit Time	Technique	Band	Threshold
L8	16 days	Dynamic Surface Water Extent (DSWE)	Multiple	N/A
		Modified Normalized Difference Water Index (MNDWI)	Green band, short-wave infrared	0.3
S1	10 days	Backscattering thresholding	–	< -13 db

S2	5 days	Dynamic Surface Water Extent (DSWE)	Multiple	N/A

In recent years, studies have shown that multi-sensor approaches combining optical and SAR data to measure inundation extent are often more robust (Ahmad, Hossain, Eldardiry, et al., 2020). Researchers have come up with various indices like Modified Normalized Difference Water Index (MNDWI) (Xu, 2006), Normalized Difference Water Index (NDWI) (McFeeters, 1996), and techniques like Dynamic Surface Water Extent (DSWE) (Jones, 2019) and angle looking SAR. Optical satellites like the Landsat series (Acharya et al., 2016; Frazier & Page, 2000; Jones, 2019) MODIS sensors onboard the National Aeronautics and Space Administration (NASA) Terra and Aqua satellites (Brakenridge & Anderson, 2006), and Visible Infrared Imaging Radiometer Suite onboard Suomi National Polar-orbiting Partnership (NPP VIIRS) (C. Huang et al., 2015) can be used to study water surface area and volume of water stored. However, a major drawback of optical satellites is that they cannot penetrate clouds. To overcome the issue of cloud cover, synthetic aperture radar (SAR) can be used with an understanding of the proper threshold on backscattering to detect water surfaces (Liang & Liu, 2020). However, SAR may not always be accurate because other smooth surfaces and shadowed areas share almost identical scattering properties with water surfaces. For example, bare soils can sometimes create false-positive cases (Shen et al., 2019). Despite such a wide range of available techniques, the uncertainty of surface water area and hence volume estimation due to the choice of methods has not been rigorously studied for lakes and wetlands. Understanding these uncertainties is challenging yet important. It is challenging due to cloud cover and seasonally contrasting environments. For example, freezing/thawing of lakes in higher latitudes can make detection of variations in volume impossible (Mayr et al., 2021; Shen et al., 2019). Similarly, lake area cannot be regularly detected due to extensive cloud cover, for example during months-long monsoon seasons.

Both optical and microwave angle-looking sensors can only estimate the area of the water bodies. On the other hand, satellite altimeters such as Jason 3, Sentinel 3, SARAL/AltiKa, provide water surface elevation . (Baup et al., 2014; Crétaux & Birkett, 2006) developed three independent approaches to estimate the lake: volume high-resolution image-based volume, altimetry-based volume, and altimetry and high resolution-based volume changes. (Duan & Bastiaanssen, 2013) and (Crétaux et al., 2016) have used a combination of lake extent and water level at different dates in order to build hypsometry relationship which was then used to calculate lake extent and level simultaneously using satellite altimetry measurements. The uncertainty in elevations from altimeters can vary from a few centimeters for large water bodies to tens of centimeters for small water bodies (Crétaux et al., 2018). The limitation of altimeters is the limited spatial sampling due to the narrow width of the sampling track. On the other hand, lidar missions with very high spatial coverage like IceSat-1 or IceSat-2, have the proven potential to measure water level at very high accuracy over a large number of lakes worldwide (Cooley et al., 2021) due to their long revisit times that however lead to missing sub-seasonal variabilities and rapid changes of lake levels.

To overcome the combined challenges of the current fleet of sensors and the limitations of existing in situ gauge networks, one possible solution to monitoring lake water level is the application of citizen science in monitoring waterbodies (Ahmad, Hossain, Pavelsky, et al., 2020; Little et al., 2021; Lowry & Fienen, 2013). Citizen science is an emerging science where the public participates and collaborates in scientific research to increase knowledge. One example of the use of citizen science is the Lake Observation by Citizen Scientists and Satellites (LOCSS) (https://www.locss.org/) program, where citizen scientists report the water height elevation of lakes or wetlands by reading staff gauges (Little et al., 2021). Hereafter, we use the terms height and elevation interchangeably to refer essentially to the vertical dimension of lakes reported by citizen scientists to estimate volume change. The objective of the LOCSS project is to work with stakeholders and local communities, who are responsible for understanding and documenting the physical behavior of lakes or depend on lake information for decision-making activities. The purpose of this study is to understand how different methods for estimating lake volume change combining satellite measurements of inundation extent with LOCSS measurements of water surface elevation, impact our ability to accurately detect variations in lake volume. By exploring an ensemble of methods and sensors to estimate area and consequently volume changes, we can derive a robust understanding of estimation uncertainty for lake volume changes. This understanding can be further nuanced for a given region that is unique to the season and other geophysical drivers such as cloud cover, rainfall, topography and water surface temperature in regions where lakes freeze.

Our study explores uncertainty in volume estimation, which can provide valuable information to the decision-makers or stakeholders to make more robust decisions based on estimation uncertainty. There are various factors that affect the uncertainty in volume estimation. For example, in South Asia, a key source of uncertainty is likely to be cloud cover during monsoon for the optical sensors and inundated vegetation for SAR microwave sensors. At higher latitudes or mountainous regions where lakes freeze, the area estimation may be more challenging due to the limitations of detecting inundation variations due to ice cover or due to the shadow effect of the high topography.

2.3. Study Site

To better understand the complex nature of uncertainty in volume estimation and how it varies at different locations, we monitored 106 lakes and wetlands globally from the LOCSS program (Figure 2.2). We focused on water bodies from the South Asian region (Bangladesh, Nepal and India). For North America, LOCSS lake height data was obtained from water bodies located in Illinois, Massachusetts, New Hampshire, North Carolina, New York and Washington

(Figure 2.3). In Europe, we had LOCSS lake height data located in South of France in the Pyrénées mountain.

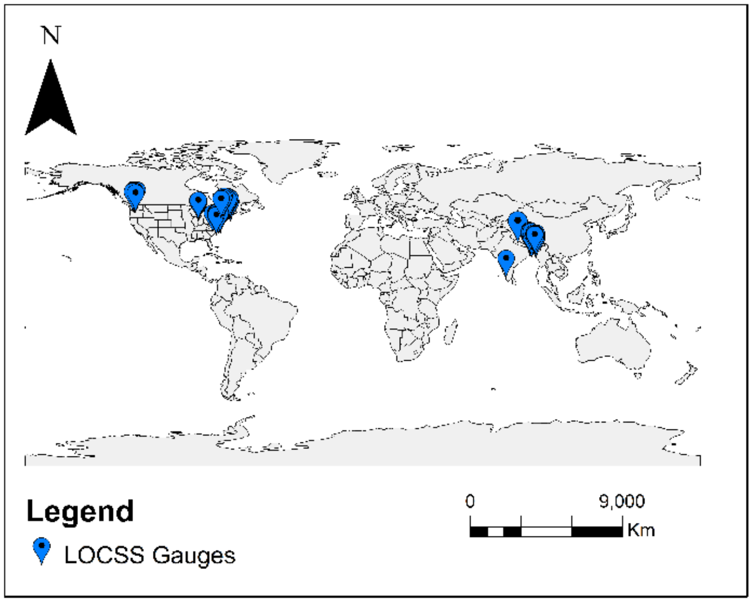

Figure 2.2: Location of LOCSS sites for the citizen science monitoring of lakes and wetlands.

The lakes in the USA and France are perennial, with some that freeze during winter. On the other hand, most of the lakes and wetlands in Bangladesh are seasonal, where water accumulates during the months of the monsoon (May to November).

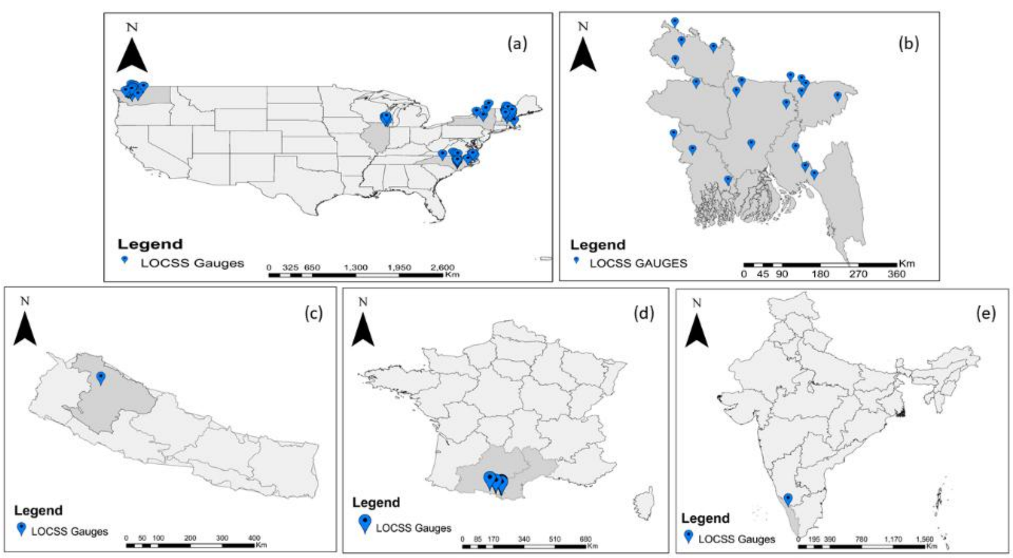

Figure 2.3: Location of LOCSS gauges in a) USA (71 lakes), b) Bangladesh (20 lakes), c) Nepal (1 lake), d) France (13 lakes), e) India (1 lake)

2.4. Dataset

For estimating the surface water area, satellites missions Sentinel 1, Sentinel 2 and Landsat 8 were used. Sentinel 1 has C-band synthetic aperture radar imaging that can penetrate the clouds and has a spatial resolution of 10 m and revisit time of 6 days. Imagery from the Sentinel 2 Multispectral Instrument (MSI) were used with a spatial resolution of 10 m and revisit time of 5 days. Landsat 8 Operational Land Imager (OLI) Tier-1 Surface Reflectance with a spatial resolution of 30m and revisit time of 16 days was used. These sensors were chosen as they were publicly available and have shown skill in detecting water surfaces (Z. Du et al., 2014; W. Huang et al., 2018; Yang et al., 2017; Yao et al., 2019). The satellite data are freely available on Google Earth Engine, a cloud-based computing platform ideally suited for a global study of lakes (Gorelick et al., 2017).

The water elevation data were collected from the citizen scientists engaged or partnered via the LOCSS program. For example, lake water height data from South Asia were obtained from citizens engaged with the relevant state or national government water agencies, such as Bangladesh Water Development Board for Bangladesh, Kerala Centre for Water Resources Development and Management for India and Nepal Department of Hydrology and Meteorology for Nepal. Similarly, most lake height data over the US were obtained from citizen scientists in the area with gauges maintained by local partnering organizations. In France, the lake heights were collected by hikers who had sent photos of the gauges via smartphone. For more details, the reader is referred to Little et al. (2021) and www.locss.org. A previous study on LOCSS has shown the water elevation data from citizen scientists are reliable and accurate when compared to automated gauges (Little et al., 2021). Nevertheless, all LOCSS data were subject to a quality control to filter out human errors that represented clear outliers. A clear outlier is one where the lake water height data is found to be a random anomaly from the underlying trend observed before and after. Such outliers were replaced with a 95% percentile threshold shown in Figure 2.4 below. For the case of France, some photos sent may also be grainy and not readable. The presence of such outliers occurred in less than 0.1% of the data. LOCSS gauges were installed in 2017 in the USA and France, 2019 in Bangladesh, 2021 in Nepal.

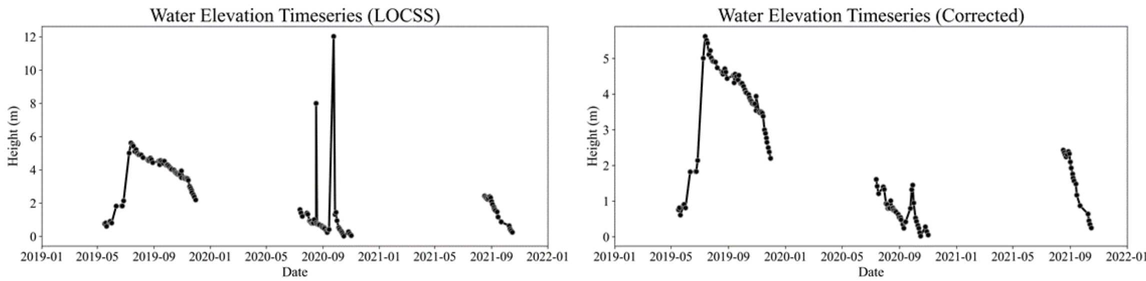

Figure 2.4: Example of water surface elevation before and after correction of outliers

2.5. Methodology

The flowchart for the methodology followed is shown in Figure 2.5. The methodology has four key components as follows: i. extracting water surface area of lakes and wetlands; ii. estimating the volume stored for all the water bodies; iii. repeating steps i) and ii) using other methods (see Table 1.1) to create an ensemble of estimates; iv. comparing the uncertainty in estimated volume as a function of region, nominal lake area and geophysical factors such as cloud cover, precipitation and water surface temperature. From here onwards we will use uncertainty and uncertainty in volume estimate interchangeably.

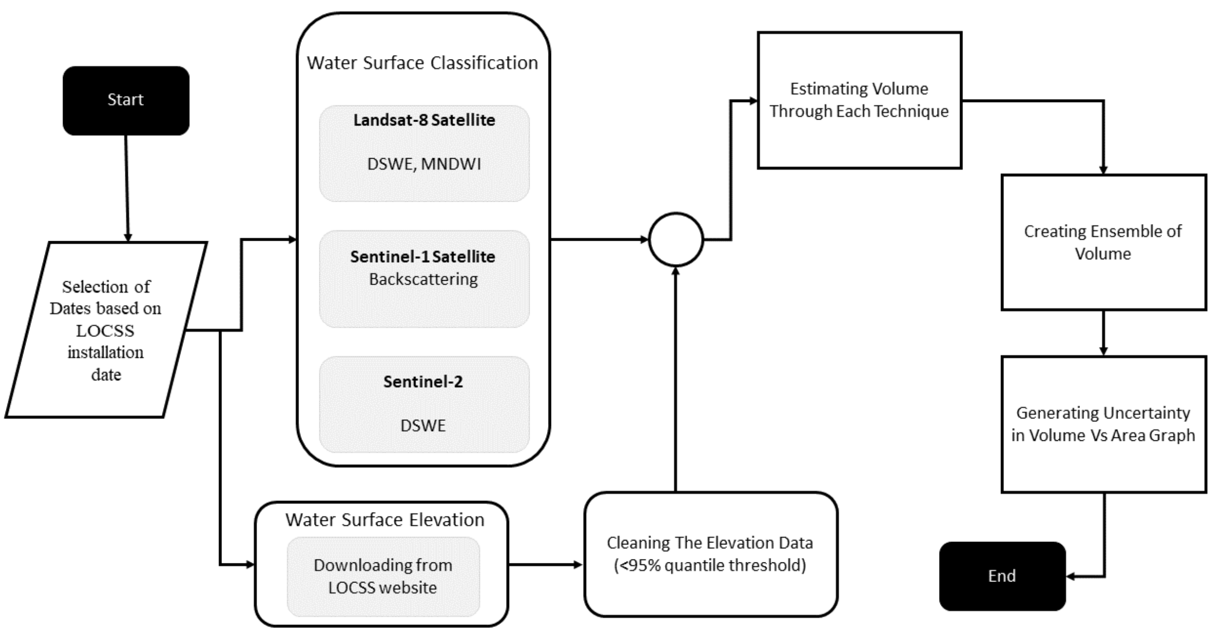

Figure 2.5: Flowchart for methodology used for exploring uncertainty of satellite-based lake volume estimation

2.5.1. Extracting Water Surface Area

2.5.1.1. Landsat 8

The Landsat 8 OLI/TIRS (L8) sensor was used to estimate the water surface area through a variety of water classification techniques. Atmospherically corrected L8 data using the (Land Surface Reflectance Code) LaSRC (Vermote et al., 2018) was used for the study. Two water classification techniques were used. The first was the Dynamic Surface Water Extraction (DSWE, (Jones, 2019)). DSWE has the ability to extract the water surface where the pixel is partially covered with vegetation and water. In addition to Landsat imagery, DSWE uses a digital elevation model, slope, hill and cloud shade. These parameters are calculated using the Fmask function (Zhu & Woodcock, 2012). The output of the DSWE consists of six possible classes: Not water, water - high confidence, water - moderate confidence, potential - wetland/ partial surface water conservative, and masked out due to the cloud, cloud shadow, or snow. The second technique used to extract the water surface area is the Modified Normalized Difference Water Index (MNDWI). (Xu, 2006) developed the definition using the NIR band with (Short Wave Infrared) SWIR band to detect the water feature in built-up areas where a threshold of 0.3 for the MNDWI was found to be a robust choice (Acharya et al., 2018; H. K. Zhang et al., 2018).

2.5.1.2. Sentinel 2

Optical imagery from Sentinel 2 (S2) sensor has a spatial resolution of 10 m, which is an improvement over the Landsat 8 spatial resolution of 30m. The DSWE technique was also applied to Sentinel 2 images. As the DSWE algorithm was designed specifically for the L8 images, scaling of S2 reflectance data is required to make DSWE work for S2 data. Surface reflectance transformation functions between S2 and L8 can be used to transform the S2 bands to L8 bands. In the study, we used the transformation function developed by (H. K. Zhang et al., 2018) to linearly map the S2 bands to L8 bands and use the DSWE algorithm. For the MNDWI technique on S2 imagery, no transformation is required according to the study conducted by (Y. Du et al., 2016).

2.5.1.3. Sentinel 1

Sentinel 1 is an angle looking C-band SAR that sends radar signals which can penetrate clouds. Water classification using the Sentinel 1 imagery was accomplished with the help of the backscattering thresholding technique. Non-water surfaces usually have high roughness and thus, they have high backscattered energy as compared to the water-like surface. The water-like surfaces appear dark in the imagery because of their smooth surface. Hence, this phenomenon can be used to extract the water surface extent by putting a threshold on the backscatter values. However, one of the drawbacks of the SAR is speckle noise, which degrades the quality of the image and causes information loss. Over the years, various techniques have been used to reduce the speckle noise such as wavelet transform (Choi &

Jeong, 2019) mean-median filters (Klogo et al., 2013). We used a focal median filter with a 30 m × 30 m window. Incidence angle also plays an important role in the image pre-processing; for the water surface classification, we considered look angles from 31.7º to 45.4º. More details on this choice are described by (Ahmad, Hossain, Pavelsky, et al., 2020). With the pre-processed image, a backscatter threshold of -13db was selected to identify the water body, as suggested by (C. Liu, 2016).

2.5.2. Extracting Water Surface Elevation

The water surface elevations were gathered with the help of citizen scientists. The data for all the water bodies were downloaded from the LOCSS website where they are publicly available.

2.5.3. Estimating The Volume Stored and Generating Uncertainty Ensemble

After estimating the water surface area and extracting the water surface elevation, volume of water stored above the minimum observed level was estimated for each of the techniques and sensors. Satellite water extent data were used for days that matched or were within 3 days of the measurement date of citizen scientists from LOCSS. To estimate the volume variation, we linearly interpolated the water surface area data, so that the timestamps of both water elevation and water surface area are the same, which makes it easy to calculate the volume change. The information of the exact bathymetry of the water bodies was not available, so we estimated the volume stored with respect to the lowest observed water surface elevation in the time series, similar to (Ahmad, Hossain, Pavelsky, et al., 2020) who had earlier applied citizen scientist height data for northeastern wetlands of Bangladesh. The lowest water elevation observation was obtained from the LOCSS website over the period of the study. For simplicity, many studies in the past have assumed trapezoidal bathymetry (Ahmad, Hossain, Pavelsky, et al., 2020; Little et al., 2021), so we also assumed trapezoidal bathymetry as shown in Figure 2.6.

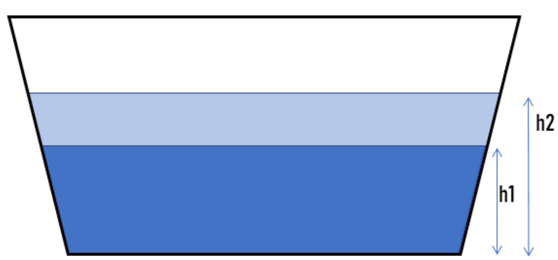

Figure 2.6: Trapezoidal Bathymetry

Pyramidal bathymetry of lakes can also be assumed as proposed in (Crétaux et al., 2016) but internal comparisons done between both hypotheses have usually yielded very similar results. In Figure 2.6, assume h1 and h2 are water surface elevations reported by citizen scientists at different instances of time. We then apply the corrections in water surface elevation (if needed) and try to estimate the water surface area A1 and A2 at the elevation reported timestamps through the satellite as shown in Figure 2.7.

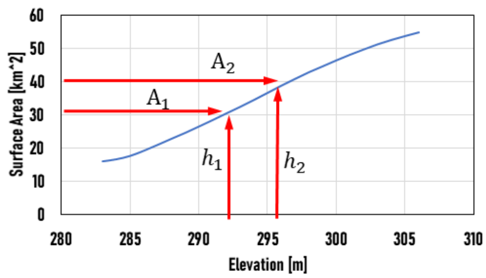

Figure 2.7: Getting Satellite Imagery on at nearest timestamp of water surface elevation

The volume of water stored between two such timestamps is estimated as per trapezoidal rule as shown in Equation (2.1.

$$\Delta S = A_{avg.} * \Delta \mathrm{h} = \frac{(A_2 + A_1)}{2} * (\mathrm{h}_2 - \mathrm{h}_1)$$

(2.1)

Hence volume stored by the water body at a given time can be calculated as:

$$V_t = \frac{(h_t - h_{min})(A_t + A_{min})}{2} \qquad \text{[L3]}$$

(2.2)

Here in Equation (2.2, h_t is the water elevation at time t, h_{min} is the lowest water elevation of the time series at each lake. A_t is the area of the lake at time t and A_{min} is the minimum area of the lake. The volume estimated in this fashion using Equation (2.2 yields the volume that can be estimated from the lowest level observed in the satellite record. Understandably, this approach may yield large errors when the difference between h_t and h_{min} is large enough to disqualify the assumption of trapezoidal bathymetry between those two heights. In our scrutiny of bathymetries above the minimum observed level, lakes that experience large height differences of many meters, such as in Bangladesh (South Asia), follow a very flat and steady trapezoidal bathymetry. In regions where bathymetry shape may be irregular over large heights, such as in the studied lakes of Europe, USA, India and Nepal, the height differences reported by citizens are usually not large enough. Figure 2.8 and Figure 2.9 shows the topography of the surrounding area of the water body and the water surface elevation.

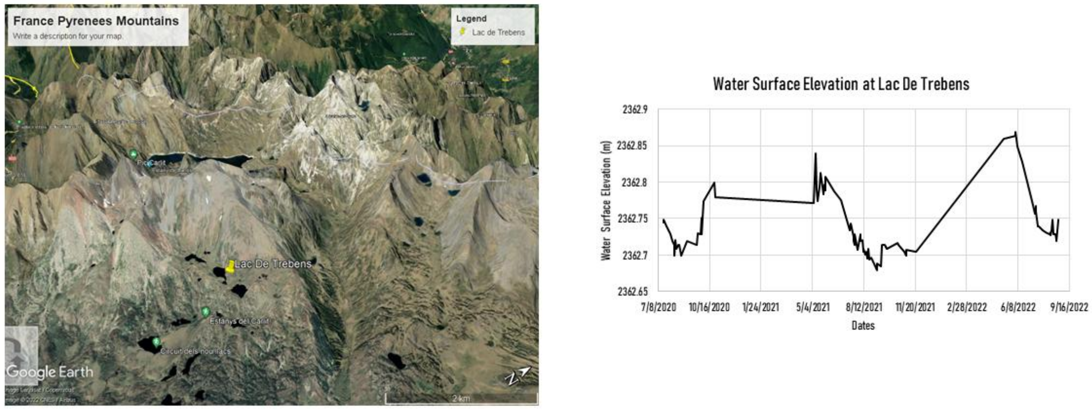

Figure 2.8: Topography and Water Surface Elevation of a Lake in France

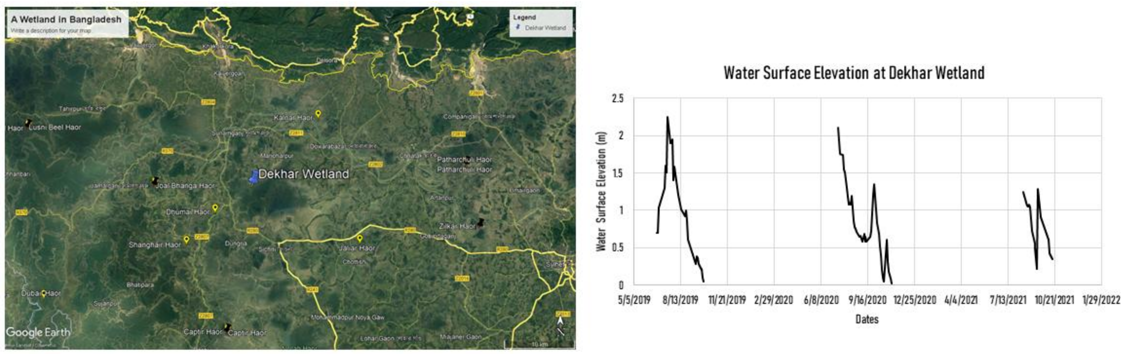

Figure 2.9: Topography and Water Surface Elevation of a Wetland in Bangladesh

The volume stored was estimated for all four techniques used in the study (Table 2.1), and an ensemble of the volume estimates was generated. Figure 2.10 shows an example of the ensemble of estimated volumes.

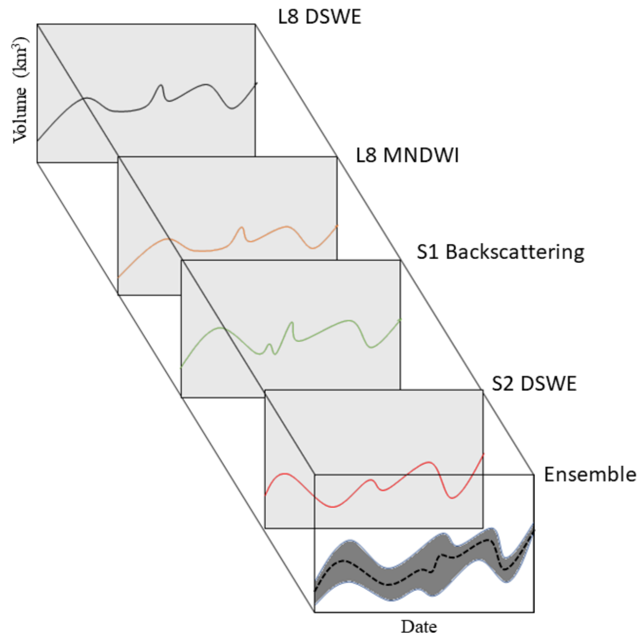

Figure 2.10: Generating the ensemble of estimates on volume stored

2.5.4. Studied Factors Affecting Uncertainty

2.5.4.1. South Asia

To understand the complexity of uncertainty in estimating volume, various factors contributing to the uncertainty were studied. Countries such as Bangladesh, India, and Nepal have a monsoonal climate, which brings extensive cloud cover and a high amount of rainfall for 3-5 months. Hence precipitation patterns and cloud cover were compared with uncertainty. Optical sensors have a limitation that they cannot penetrate the clouds. The complementary nature of optical and radar sensors with unique strengths and weaknesses collectively give rise to estimation uncertainty. Gridded precipitation data for Bangladesh was downloaded from the ERA5 hourly precipitation and Bangladesh Meteorological Department data library and gridded precipitation data for India was downloaded from Indian Meteorological Department. The cloud cover data was collected from information provided in the Landsat 8 satellite data product.

2.5.4.2. North America

In the regions of North America studied here, the monsoon is not as dominant, unlike South Asia. We therefore studied the uncertainty in volume estimation as a function of temperature and cloud cover. The water surface temperature of lakes was estimated using the Landsat 7 Collection 1 Tier 1 (L7). Low-gain Thermal Infrared 1 Band (B6_VCID_1) was used to estimate the temperature. While the cloud cover over the water bodies was estimated using L8.

2.5.4.3. Europe

In Europe, we studied the water bodies in France. All of the lakes in France were in highly mountainous areas of the Pyrenees. Thus, with the exception of Nepal, these lakes were located in the most topographically variable landscape of any LOCSS lakes. These lakes also freeze in the winter. We ran the same analysis on France as we did on North America.

2.5.5. Estimating The Uncertainty In Volume Estimation

We chose a metric for uncertainty in volume estimation that provides us with an idea for average spread of the estimated volumes over time relative to the statistically expected volume of a water body (also over time). Here the expected volume is assumed to be the arithmetic mean of the volumes estimated by the four methods. We call this metric the "time-averaged uncertainty metric." This time-averaged uncertainty metric is calculated using Equation (2.3). Here we use the time-averaged uncertainty metric in relative terms normalized by the mean volume, to allow comparison across all lakes and regions. A time-averaged uncertainty metric value of less than one means that the current suite of sensors and methods is generally able to estimate volume variations with a spread that is less than the mean value, and hence the uncertainty may be considered acceptable at any given time. Vice versa, an uncertainty metric value of more than 1 means the spread of uncertainty is significantly larger than the mean value itself, and hence the volume uncertainty may be considered unacceptable at any time. Figure 2.11: Volume stored and Uncertainty time-series for Korchar wetland in Bangladesh. illustrates how the time-specific uncertainty in the volume varies for the Korchar wetland in Bangladesh over time to yield the time-averaged uncertainty metric defined in Equation (2.3).

$$Time\ averaged\ Uncertainty\ = \frac{\sum_0^n \left[\frac{Max\ Volume_t - Min\ Volume_t}{Mean\ Volume_t}\right]}{\sum_0^n t} \tag{2.3}$$

The time-specific behavior of uncertainty is particularly suited for developing a temporal understanding of seasonal water bodies, such as wetlands in South Asia. During the development of a wetland in the monsoon season, the spread of the ensemble may be smaller yet the uncertainty metric for that specific time can be higher because of time-specific low mean for estimated volumes. Such a high time-specific uncertainty can be indicative of the limitation of the sensors for water bodies with very small volumes and variations at that time. As these wetlands develop and the volume stored increases, the time-specific uncertainty metric can decrease if the collective precision of the sensors hold. Conversely, the opposite can happen with time-specific uncertainty rising as volume increases. We show one such example in Figure 2.11 (a) and (b). A red line is shown to demonstrate the case for a wetland in Bangladesh where the time specific uncertainty rises despite increase in volume after the height of the monsoon in August. This corroborates the fact that uncertainty of volume estimation can be dependent on many factors, many of which are time-specific (such as cloud cover, land temperature, growth of vegetation, and irregular/non-trapezoidal bathymetry).

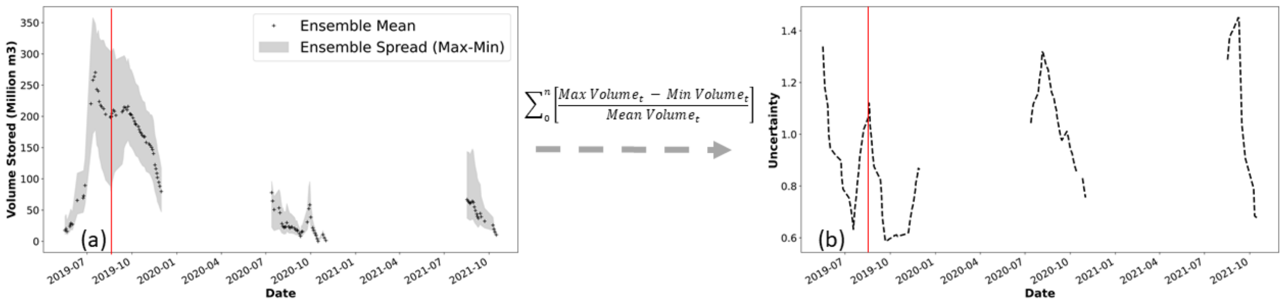

Figure 2.11: Volume stored and Uncertainty time-series for Korchar wetland in Bangladesh.

2.6. Results

In this chapter we demonstrate a few examples of time varying uncertainty (not the time averaged uncertainty) that are representative of lakes and wetlands for their regions. Figure 2.12 shows the volume estimation uncertainty of a wetland in Bangladesh. In general, the wetlands in Bangladesh are fully inundated during May to December, while and from January to April, they are often dry. It is seen that during higher cloud cover and precipitation, the uncertainty spread is high. Figure 2.13 shows the ensemble of Pookode lake in Kerala, India. Kerala in general receives two monsoons. One is the southwest monsoon (June - September) and the other is the northeast monsoon (October to December). Essentially, the entire period of June to December is characterized by extensive cloud cover. We observe that volume stored and uncertainty in volume are both higher as the monsoons retreat in December with gradual decrease as cloud cover and precipitation decreases in April. The pattern repeats itself from June to December again as the two monsoon seasons complete their cycle.

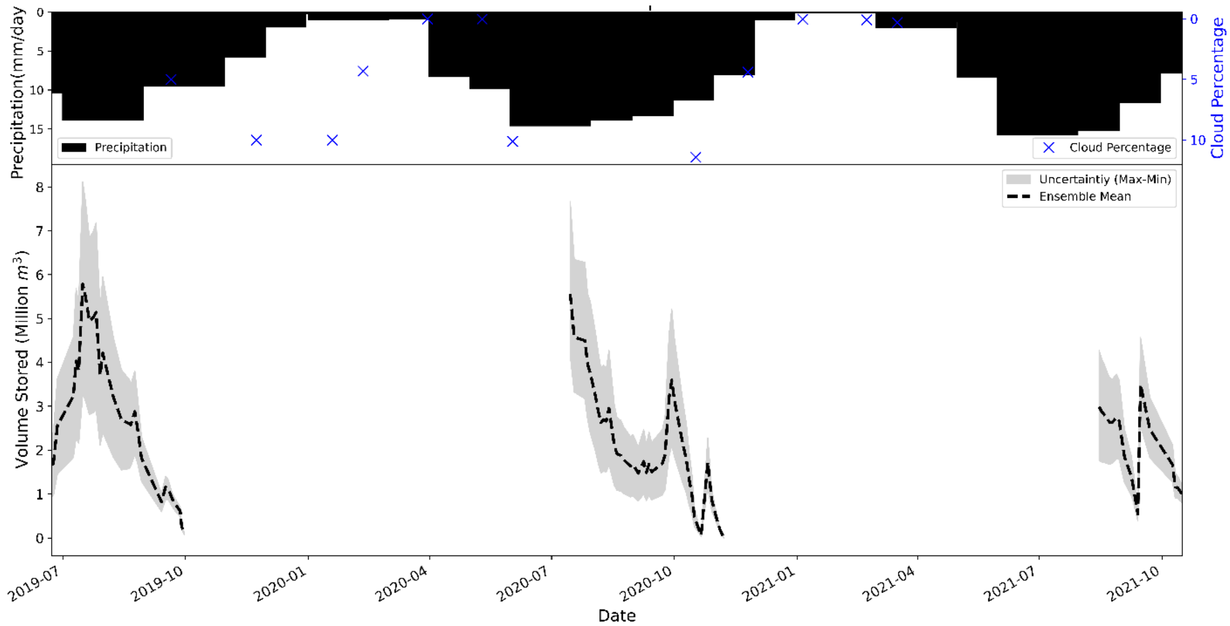

Figure 2.12: Time series of uncertainty of volume estimation of Dekhar Haor wetland in Bangladesh

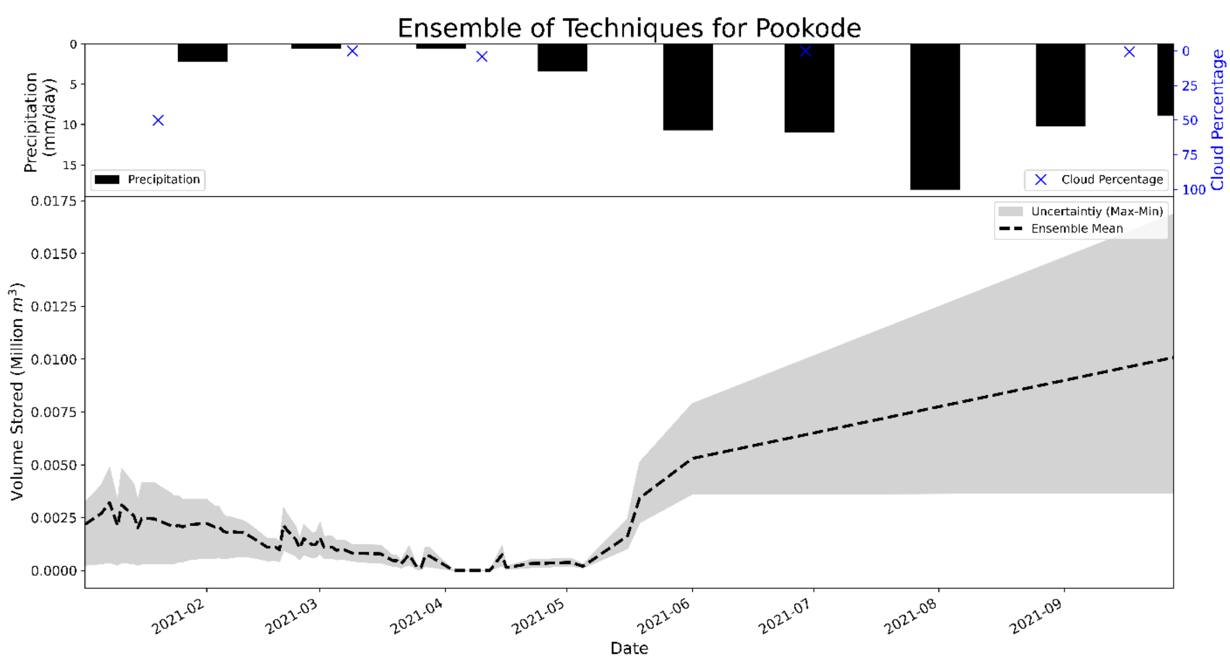

Figure 2.13: Time series of uncertainty of volume estimation in Pookode lake in Kerala (India).

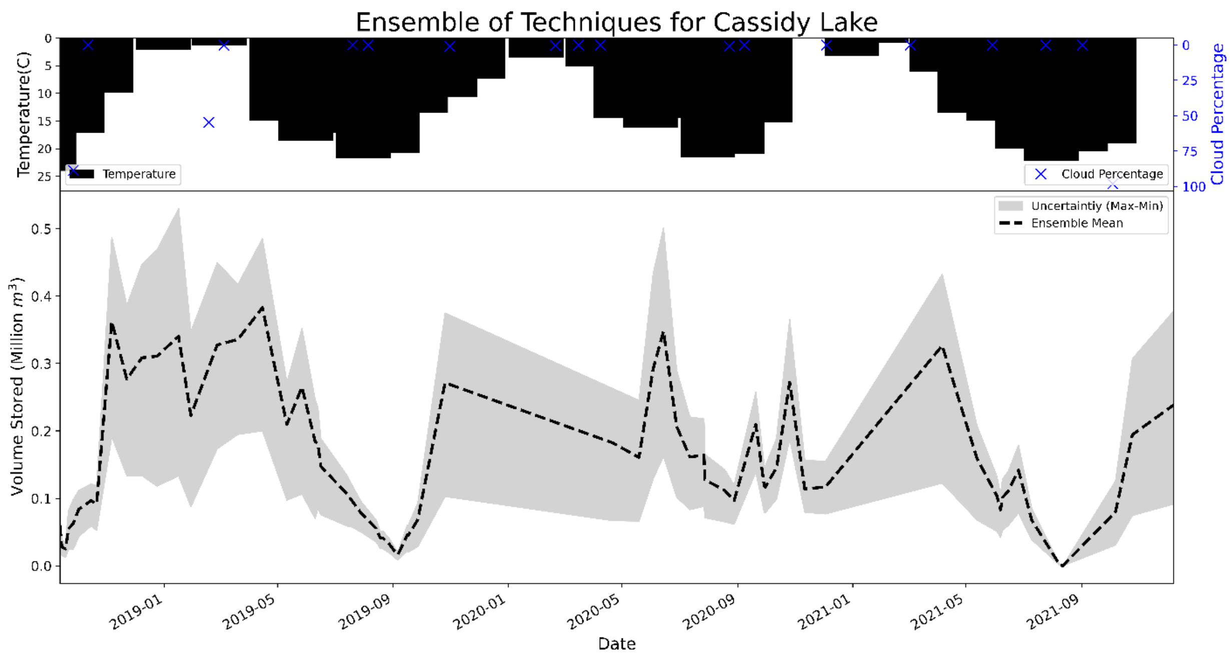

Figure 2.14: Ensemble of Cassidy Lake in Washington state.

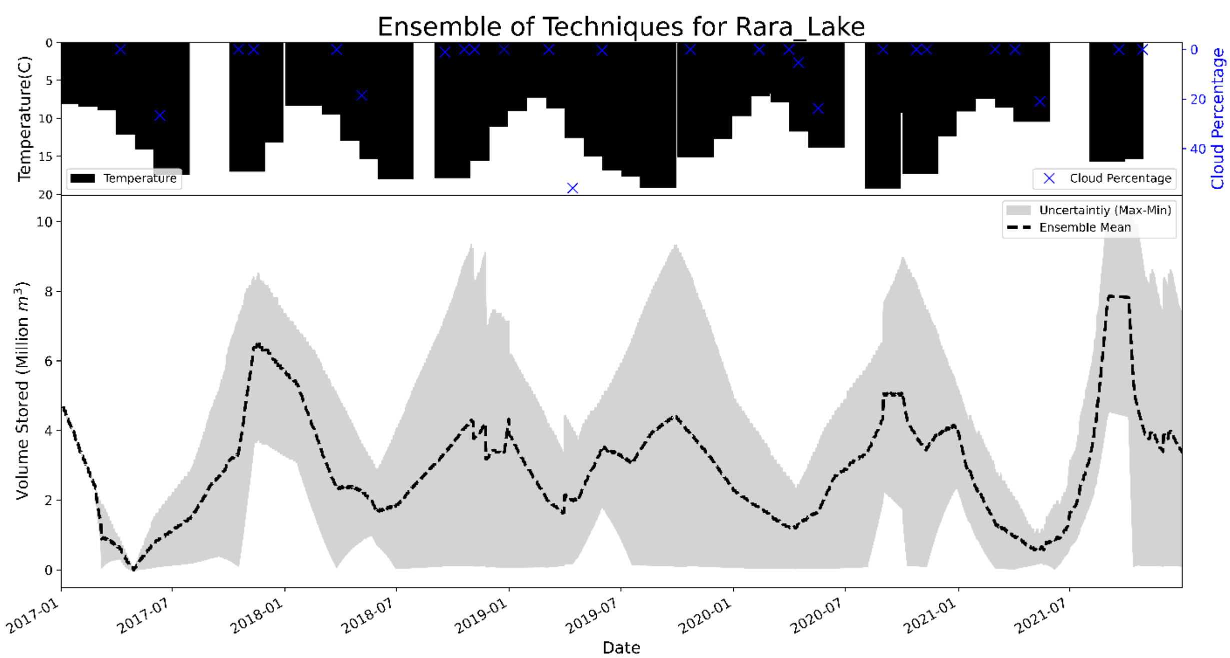

Figure 2.15: Ensemble of Rara Lake in Nepal.

For US lakes, we studied volume uncertainty as a function of cloud cover and water surface temperature given the tendency of some lakes in upper latitudes to freeze during winter. In Figure 2.14, we show the uncertainty spread for Cassidy Lake (Washington State), which is found to be high during freezing conditions. When volume stored is low, the uncertainty spread is also found to be quite high. As there are likely many other controlling factors, water temperature provides only a partial explanation of the temporal behavior of uncertainty.

Figure 2.16 shows the time-averaged uncertainty metric (Equation (2.3)) for all the water bodies as a function of the nominal lake area. This is a plot showing the aggregate behavior volume estimation uncertainty for each region as lake area changes. The idea is to understand if there is a threshold area for a lake size below which the time averaged uncertainty metric is high (> 1). From Figure 2.16 the percentage of water bodies having uncertainty metric less than 1 in Washington, South East Asia (Bangladesh, India and Nepal), Illinois, East Coast USA (New York, Massachusetts, North Carolina) and France was found to be 56%, 45%, 75%, 71% and 62% respectively.

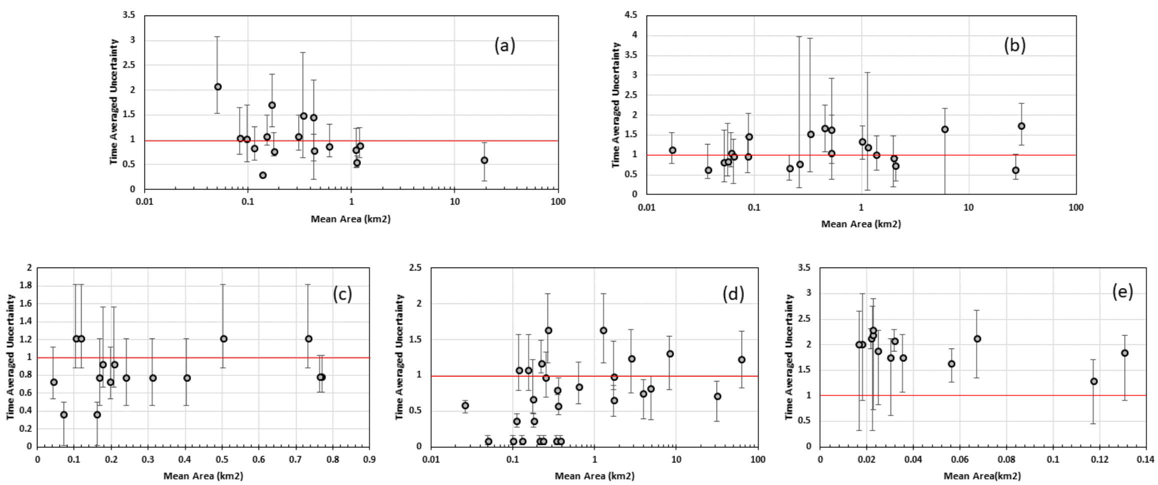

Figure 2.16: Time-averaged Uncertainty in volume Vs Nominal lake area for- (a): Washington (USA), (b): Bangladesh, India and Nepal, (c): Illinois, (d): New York, Massachusetts, North Carolina, (e): France

To understand the role played by individual area estimation methods in volume estimation uncertainty, we ranked each of the four methods from highest to lowest average volume estimates for a given lake. In Figure 2.17, we show in a four panel plot the methods for each lake with highest volume (uppermost panel), second-highest volume (middle panel), second-lowest volume (2nd panel from bottom) and lowest volume (bottommost panel). The idea is to see if performance of methods is consistent across lakes or if other geophysical factors pertaining to the lake and the ambient environment control the tendency to estimate the highest or lowest value in the ensemble. In general, the DSWE method using Sentinel-2 and backscattering method using Sentinel-1both have a tendency to yield higher volume estimates. However, when looked at as a whole, there does not seem to be a single method that is found to consistently estimate the highest, lowest volume or median volume. This indicates that in our study of lake volume estimation uncertainty, there is no single method that can be filtered out to minimize uncertainty and that all methods should be considered collectively to improve our understanding of uncertainty.

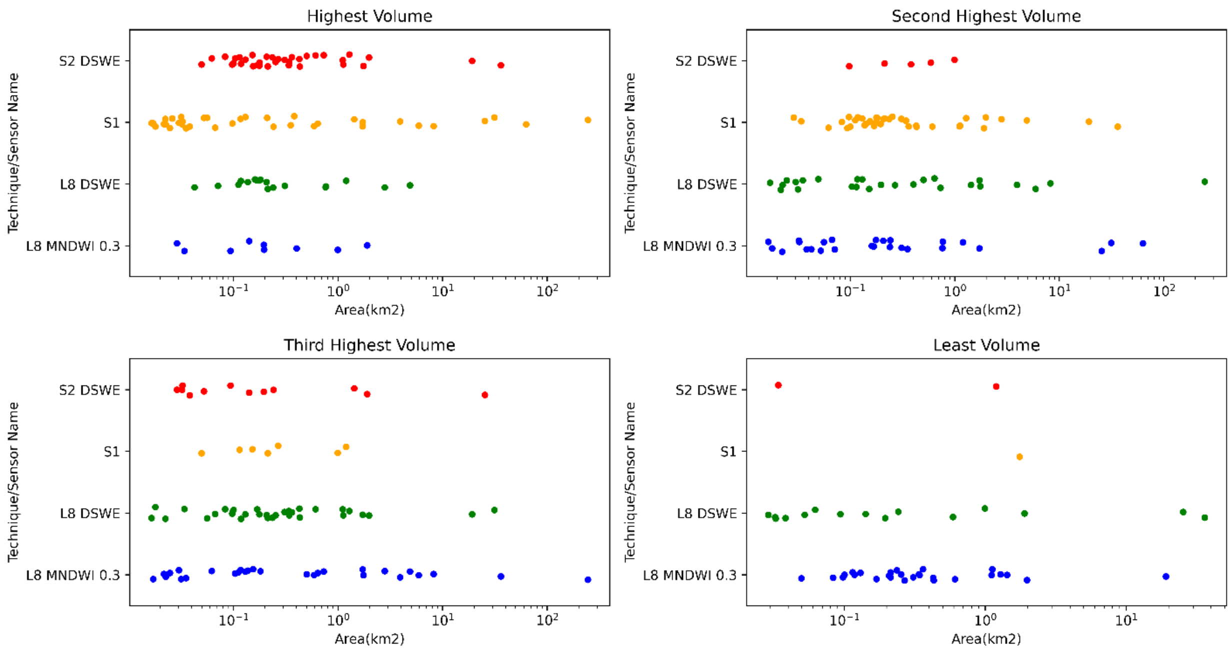

Figure 2.17: Ranking of the four methods

2.7. Conclusion

We studied 106 lakes and wetlands around the world where LOCSS gauges were installed to record water elevations measured by citizen scientists. We define a time-averaged uncertainty metric where the threshold value 1 will be used as the cutoff for acceptable uncertainty (< 1) or unacceptable uncertainty (> 1). When considering all the lakes studied, there is no clear pattern in our findings where lakes larger than a certain threshold can claim to experience higher skill in estimation of volume. However, at individual regions, there are some nuanced patterns. For example, lakes in Washington and France show a clear dependency of uncertainty as a function of area where the time-averaged uncertainty metric decreases as nominal lake area increases. In South Asia, lakes larger than 0.5 km2 experience an uncertainty metric of less than 1 near-consistently. This trend is driven mostly by lakes in Bangladesh where lakes larger than 10 km2 yield acceptable uncertainty. This implies that the flat terrain nature of Bangladesh topography combined with more dynamic hydro-meteorological and land use patterns compared to other regions studied pose a significant challenge to lake volume estimation. In the US, what is interesting is that lakes east of the 108th meridian (Colorado Rockies) exhibit considerably lower uncertainty in volume estimation. This uncertainty decreases gradually for lakes located further eastwards, starting from Illinois to Eastern US (Massachusetts, New York and North Carolina). For example, in Washington state, the average time-averaged uncertainty appears to be around 20% higher than lakes in Illinois which are about 50% higher than lakes in the eastern US. It is clear that much smaller-sized lakes in the eastern US can be estimated with considerably less uncertainty. In France, we saw that the spread of uncertainty is high. One of the plausible reasons for this can be the shadow of the mountains. The LOCSS gauges are installed in the South of France, near the Pyrenees mountains. While digitizing the lakes in France, mountains projecting the shadow on water bodies were observed. (Ji et al., 2015) discussed how mountain shadows can be misclassified as water pixels. We should however exercise caution in interpreting the volume estimation uncertainty pattern for each region (e.g., USA, France and South Asia) given that samples of lakes studied here may not necessarily be a statistically representative sample of all lakes in those regions.

Our lake height data were obtained from the citizen science program of LOCSS, which has the additional objective of validating and improving lake products anticipated from the planned Surface Water and Ocean Topography (SWOT) mission. The SWOT satellite mission is a joint mission of the National Aeronautics and Space Administration (NASA) and Centre National d'Etudes Spatiales (CNES) with contributions from the Canada Space Agency and the United Kingdom Space Agency. SWOT is planned for launch in November 2022 (Biancamaria et al., 2016). It will be the first satellite of its kind that will report water surface elevation and water surface area simultaneously with a revisit time of 21 days or less at a given location. The primary instrument on SWOT is Ka-band Radar Interferometer (KaRIn), which uses radar interferometry and synthetic-aperture radar (SAR), which gives high-resolution water elevation and inundation extent (Biancamaria et al., 2016). Currently, as noted in this study, to estimate the volume, water surface area is derived from satellite sensors while the elevations are obtained either from concurrently-flying altimeters or from the in-situ data. SWOT, with its simultaneous measurement of area and elevation, will improve our ability to estimate volume more consistently. Moreover, SWOT is a swath interferometer which will cover the whole Earth and monitor lakes larger than 250 X 250 meters. This will be an unprecedented global view of the lake storage change dynamics at global scale.

Our findings therefore have implications for the SWOT mission. First of all, the availability of LOCSS gauge data from citizens can be expected to provide valuable validation data to compare SWOT-estimated volume changes once SWOT starts to provide lake area and elevation simultaneously. Secondly, SWOT observables could be combined with pre-SWOT satellite data to create higher frequency estimates of lake volume with lower estimation uncertainty. Armed with a general idea of what regions, specific factors and the minimum lake size matter in achieving an acceptable uncertainty, LOCSS gauges can be strategically expanded or the data quality for lake storage change can be flagged accordingly.

The estimation of uncertainty for volume is useful for practical applications at ungauged regions lacking historical records, such as sizing of surface water storage facilities or flood control structures. For example, if an urban settlement is planned in the ungauged region with no historical records, where lakes are the only source of surface water, then the freshwater storage and distribution system size would need to be based on the minimum (worst case) scenario of lake volume experienced over a sufficiently long period. Similarly, a flood protection facility in the same ungauged region would have to be designed based on the maximum (worst case) scenario of lake volume observed over a long record. The range of estimation uncertainty gleaned from an ensemble of satellite sensors and techniques facilitates such societally relevant application in the design of water management facilities at regions lacking historical in-situ records. In a previous effort based on LOCSS (Ahmad, Hossain, Pavelsky, et al., 2020), the estimation of total volume stored in northeastern Bangladesh with uncertainty has already triggered a conversation by the Bangladesh Government to exploit any excess surface water for commercial revenue-generating purposes (personal communication with Director General of Bangladesh Water Development Board).

Our study is not without limitations. One key limitation is the short period of LOCSS data for many regions, such as South Asia. Another limitation is the absence of high-resolution optical imagery at the meter scale, such as Planetscope data, that can provide a better reference for lake volume under clear sky conditions (Qayyum et al., 2020). For example, volume estimated by an optical sensor at 1 m resolution could provide a much better understanding of the expected volume. We hope these limitations can be addressed in a future study as the LOCSS data continue to grow with more participation from citizen scientists around the world.

This chapter leveraged the existing LOCSS gauge network, to answer the key question of how uncertainties in estimating the volume of surface water vary across geographical, seasonal, and environmental contexts. The findings provide an understanding of how uncertainty propagates through satellite-based volume estimation and establish a benchmark for future missions. However, while this analysis characterizes uncertainty within regions equipped with in-situ gauges, it also underscores a critical limitation: such infrastructure is not universally available, nor is it practical or financially feasible to install gauges across every water body.

As the LOCSS network continues to expand into new regions or as other stakeholders plan to increase their gauge networks, a key question arises: How *should new gauges be strategically deployed to maximize observational efficiency*? Addressing this requires a systematic framework capable of identifying the optimal number and spatial distribution of gauges necessary for robust regional volume estimation. Moreover, this framework should be adaptable to integrate both physical gauge measurements and satellite-derived virtual stations from current and forthcoming missions, thereby enhancing monitoring efficiency and validation capacity. In the next chapter (Chapter 3), we address these additional questions and issues.

Chapter 3. A Network Design Approach for Citizen Science-Satellite Monitoring of Surface Water Volume Changes in Bangladesh

3.1. Abstract

This study introduces a reproducible and integrated framework for gauge network design for tracking surface water availability in a citizen science framework. For identifying the optimal number and location of gauges, the framework uses Monte Carlo simulations. With randomized selection of gauges from an existing high-density network, the framework does not impose a single optimal solution but rather provides a range of optimal solutions, allowing water managers the freedom to select the number of gauges and locations in accordance with their specific needs. Using the methodology, the study revealed how a strategic placement of gauges can potentially improve the estimation of surface water storage variations. The study findings can be used to enhance the assessment of lake products derived from missions such as the Surface Water and Ocean Topography (SWOT).

Note: This chapter has been published in its current form as an article in *Environmental Modelling and Software* (Khan et al., 2023). The only differences are in the numbering of figures, tables, and references to other sections.

3.2. Introduction

Robust and sustainable management of surface water resources is crucial in today's world (Kramer et al., 2023). An essential part of this responsibility involves the establishment and maintenance of a comprehensive network of in-situ gauges capable of monitoring pertinent water bodies. Unfortunately, financial constraints limit the extent of gauges and result in biased gauge placement in global hydrologic monitoring networks . In general, there has been a decline in these networks globally (Burn, 1997; Mishra and Coulibaly, 2009; Pilon et al., 1996; Soares and Calijuri, 2021). While the underlying causes for this decline are varied, some of the major reasons, according to Mishra and Coulibaly (2009), are "insufficient funding, inadequate institutional frameworks, a lack of appreciation of the worth of long-term hydrological data, and in a few cases, the turmoil caused by wars and other disasters." This globally declining trend of gauges emphasizes the necessity for more efficient gauge network designs.

Significant progress has already been made in the domain of precipitation gauges, with established guidelines currently provided by bodies such as the World Meteorological Organization (WMO) (World Meteorological Organization, 2008). However, the application of these principles in the monitoring of surface water contained in lakes and wetlands remains unexplored. Given financial limitations and the fact that there are approximately six million lakes worldwide larger than one hectare or 0.01 km^2 (Messager et al., 2016), it will never be possible to monitor most lakes and open-water wetlands worldwide using on-the-ground technology. The challenge thus lies in the design of a gauge network capable of delivering reliable and comprehensive data that is reasonably representative of lakes and/or wetlands in a particular region given that there will always be limited number of gauges to install. Addressing this challenge necessitates the determination of the sensitivity of gauge density along with their strategic or optimal placement. By drawing from the principles of optimal precipitation gauge network design under limited gauging (Rodríguez-Iturbe & Mejía, 1974) to the context of lakes and wetlands, water resources management can be further enhanced in an era of growing need to monitor the planet's surface water.

In the quest to develop a gauge network design methodology that is reproducible, it is incumbent on us to first experiment on a region that is already quite dense with existing gauges. Using such a dense network, we can develop the methodology and determine the costs and benefits of removing or adding gauges and their strategic locations in the form of sensitivity analyses. The Lakes Observation by Citizen Scientists and Satellites (LOCSS) project (https://www.locss.org/), was established in 2017 to increase the understanding of relatively small lakes around the world by combining satellite data with water level data collected by citizen scientists across the world. LOCSS also participates in the assessment of the Surface Water and Ocean Topography mission (SWOT; Biancamaria et al., 2016). More details on the LOCSS project are provided by Ahmad et al. (2020), Khan et al. (2023), and Little et al., (2021). Bangladesh, the country with the largest assembly of LOCSS gauges outside the United

States, has the unique distinction of having the highest gauge density, with an average of one gauge per 2500 km^2 (50 X 50 km grid). Its gauge network, primarily concerned with monitoring the water elevation of lakes and wetlands, has already offered valuable insights into surface water volume changes (Ahmad et al., 2020; Khan et al., 2023). By integrating the data from these water elevation gauges with satellite-estimated water surface areas, we can estimate the changes in volume of water stored with increased precision.

In understanding the sensitivity of gauge density to estimated surface water volume change (or volume storage, to be used interchangeably), and the strategic placement of these gauges, Bangladesh can serve as a laboratory for developing generic network design approaches due to its high hydroclimatic variability and high gauge density. The need for a network design approach is justified in part by the assessment needs of lake products from the new SWOT satellite mission, using citizen science as a major foundation. Surface water gauge network design ensures that satellite altimetry validation includes a representative range of lakes while avoiding unnecessary duplication of resource-intensive ground monitoring.

Significant methodological advancements in the design of precipitation gauge networks since the 1960s can provide useful context for designing lake and wetland networks. One of the early studies in this field was conducted by Eagleson (1967), who delved into the physiographic dynamics of catchments to contribute to network design. Building upon this foundation, Rodríguez-Iturbe and Mejía (1974) proposed a multidimensional methodology that considered both the quantity of observations from gauges and their locations. Further evolution of these ideas was seen in the work of Bras and Rodríguez-Iturbe (1976), who used multivariate estimation theory to design optimal gauge networks. Krajewski et al., (1991) then conducted a Monte Carlo gauge network simulation study with a physically based hydrologic model to simulate overland flow and streamflow. The basin response from a network of high density of precipitation gauges was assumed to be the "ground truth" used to assess simulated results from degraded network density. Their use of Monte Carlo established a new standard for network comparison. In recent years, Wu et al., (2020) suggested an optimal number of precipitation gauges using ordinary kriging and spatial correlation approaches. With the advent of satellite data, the paradigm seems to have recently shifted. Bertini et al., (2021); Dai et al., (2017); and Morsy et al., (2021) all leveraged satellite data in their studies for optimal gauge network design, but mostly for precipitation.

The complexity of interrelationships among hydrologic variables is documented in the existing literature, as highlighted by (Mishra & Coulibaly, 2009). The intricate dynamics of these variables imply that a change in one can induce a corresponding shift in another. For instance, an event of significant precipitation can lead to an increase in the value of water information captured at a gauge compared to other locations with less precipitation but with similar density of gauges. Conversely, a high rate of evaporation in a region can contribute to a reduction in the recorded water levels at a gauged site. Despite the critical nature of these interactions, the current body of research does not present any study, to the best of our knowledge, that has considered these interplays in the design of a hydrometric network for lakes and wetlands (Keum & Coulibaly, 2017; Mishra & Coulibaly, 2009, 2014). Therefore, there is an urgent need for a comprehensive framework that incorporates these hydro-physical factors into the gauge network design process. Such a framework would be instrumental in enhancing the reliability of the gauge network, thus ensuring its ability to accurately reflect the surface hydrologic conditions of the region the gauge network is meant to capture cost-effectively.

Despite the range of methodologies on gauge network optimization, a common objective underlies these studies: the estimation of streamflow or precipitation at ungauged sites is achieved using the information at gauged sites. Although, the contributions from these studies has significantly advanced the field of surface hydrology, there is a notable gap in the application of these network design principles to lakes and wetlands.. The importance of lakes and wetlands ecosystems has been highlighted by Crétaux et al., (2016), Khan et al., (2023), and Lettenmaier et al., (2015); emphasizing a need for improved monitoring. This study extends the principles of precipitation and stream gauge network design to the realm of lakes and wetlands. Such an approach has the potential to significantly advance surface water resource management through synergistic use of satellite remote sensing and on-ground gauging and citizen science techniques. Consequently, this can enrich our understanding of resource allocation for its monitoring, and equip us to meet future environmental challenges. We aim to address two region-specific key research questions in this study:

Q.1 ***What is the optimal gauge density needed to derive surface water volume changes using citizen science-based water levels and satellite data?***

Q.2 ***How does uncertainty in surface water volume change estimation vary as a function of gauge density and hydro-physical regime characteristics?'***

Through harnessing of the high gauge density in Bangladesh and its hydroclimatic or hydro-physical variability, we aim to build an approach that can be a starting point for developing region-specific lake gauge network optimization strategies around the world. For example, given a priori gauges of reasonable density in regions where surface water requires extensive monitoring, the methodology developed here for Bangladesh should be reproducible to allow further optimization of the existing network by adding, rearranging or reducing gauges taking into account the region's unique hydro-physical regime, lake distribution and storage patterns. The findings of our study and the answers to the above questions are conditioned on the assumption that the current network of dense LOCSS gauges in Bangladesh offers a reasonable benchmark to compare against and that a higher density of lake gauges (than in Bangladesh) is practically unfeasible in most places around the world, including developed regions. Because it is fundamentally impossible to derive the true surface water storage variability no matter how dense a gauging network is, we believe that our findings and developing a reproducible methodology for optimization of a gauge network, relative to this Bangladesh benchmark, has merit in the current state of the art of very limited lake gauging around the world using on-ground technology.

3.3. Study Area

Since the start in 2017, the LOCSS project has strategically positioned numerous gauges across the globe, with the densest concentration in Bangladesh. Figure 3.1 summarizes the global dispersion of LOCSS gauges with a zoom into the specific location of gauges within Bangladesh. Notably, it is this very high gauge density, relative to other nations where LOCSS gauges have been installed around the world, that allows the development of a gauge network design framework. Furthermore, Bangladesh by virtue of being a delta with distinct dry and wet seasonality, is host to water bodies that show very diverse and dynamic range of storage patterns over time and space unlike most regions. For example, wetlands are frequently formed each year at the end of the dry season when they are empty and drained completely after the end of the previous wet season. Because of this cradle-to-grave formation of wetlands each year, many lakes undergo significant water level changes of a few meters within a year, and lake formation is driven either by precipitation, or a combination of precipitation-driven runoff and transboundary flow from mountainous regions. These inherently diverse features of water body formation and their dynamics offer Bangladesh as a test case to develop a methodology that can be reproduced for many other regions around the world. In this study, there were 59 high quality gauges dispersed in an area of about 140,000 km^2 representing a density of 1 gauge in every 50 km X 50 km grid.

Figure 3.1: Global Presence of LOCSS Gauges for lakes with a zoom into Bangladesh – the region with the densest gauge density

3.4. Network Design Framework

Figure 3.2 shows the proposed network design framework, with a brief description of each step summarized below followed by a detailed methodology:

Step 1: The identification of hydro-physical regimes is the first step. This is conducted based on climatology of precipitation and temperature, terrain roughness, and the density water bodies present within a specific region.

Step 2: Estimation of the volume of surface water that is stored in gauged water bodies using the ensemble technique (Khan et al., 2023).

Step 3: Generation of Area-Volume Curve for each hydro-physical regime, which allows estimation of total water storage across all water bodies. This is achieved by using the highest density of gauges. A Regression Line is then fitted to the Area-Volume (A-V) Curves, using the proposed methodologies.

Step 4: Repetition of steps from 1 to 3, with systematically degraded gauge density based on randomized selection of gauges per the Monte-Carlo simulations. This step is a sensitivity study to understand how gauge density (adding or removing gauges) responds to the estimation of volume changes.

Step 5: Development of the relationship between the gauge density and uncertainty (or percent error), treating the results from the highest density as the benchmark. This step provides the summary of the sensitivity of gauge density relative to the maximum density that was available.

Step 6: Development of criterion for the strategic placement of gauges. This is derived from a 'degraded' density and expected uncertainty as a function of hydro-physical similarity in surface water variability.

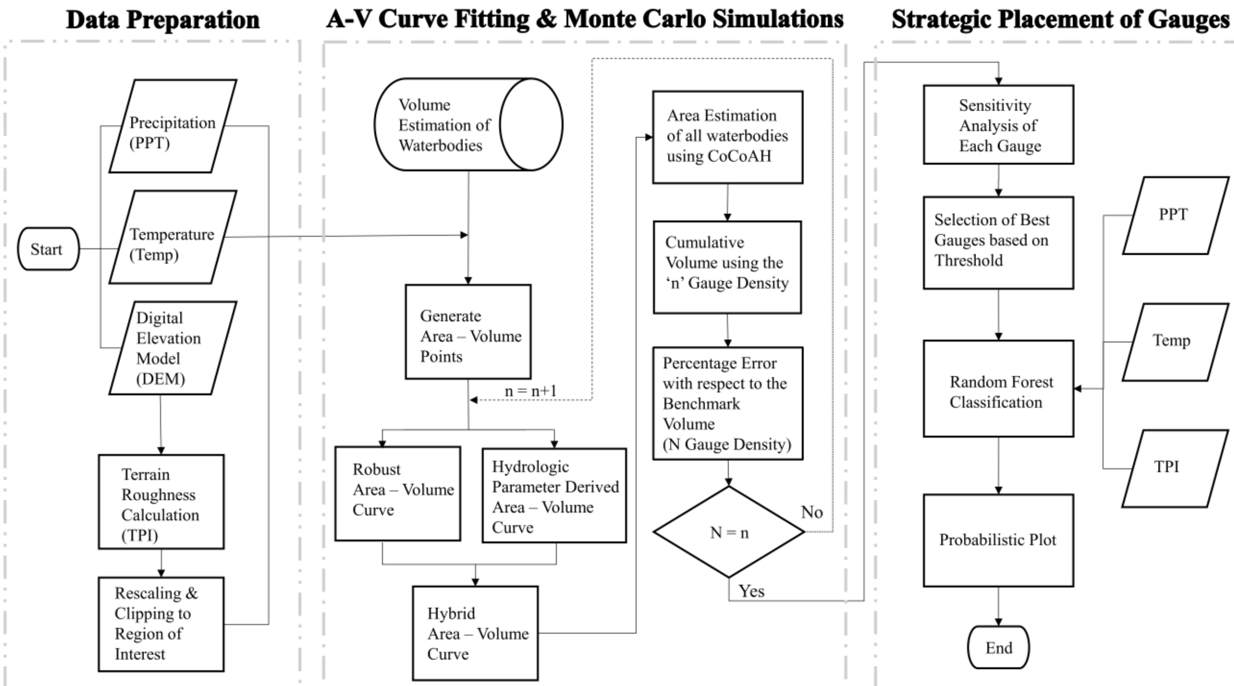

Figure 3.2: Flowchart Depicting the Proposed Methodology for Optimal Gauge Network Design for lakes/wetlands. The three types of data input shown on the far right (precipitation, temperature and terrain roughness) are derived from the steps in the far left under "Data Preparation."

3.5. Methodology

3.5.1. Data Preparation for Hydro-Physical Regime (Step 1)

The volume of surface water stored in a water body can be controlled by a number of hydro-physical factors, key among them are precipitation, evaporative losses, and topographic roughness. To facilitate our research on understanding the hydro-physical regimes, we sourced multiple datasets, focusing on the key factors, which are detailed in Table 3.1.

Table 3.1: Dataset Used in the Study

Data Type	***Source***	***Spatial Resolution***	***Range (years)***	***Citation***
Precipitation	ERA 5 Reanalysis	0.25° Grid	30 (1993-2022)	Hersbach et al., (2020)
Temperature	ERA 5 Reanalysis	0.25° Grid	30 (1993-2022)	Hersbach et al., (2020)
Digital Elevation Model	ASTER Global DEM	30 m Grid	---	NASA/METI/AIST/Japan Space systems And U.S./Japan ASTER Science Team, (2019)

The objective of estimating the hydro-physical regime is to develop a methodology for gauge network design, given the limited number of *a priori* gauges, that can be reproduced on the basis of governing features that control surface water storage variability in other regions of the world. Based on these three key parameters, Bangladesh represents multiple hydro-physical regimes. The northeastern region, locally referred to as the “Haor,” experiences heavy monsoon rainfall and subsequent flooding from May to December, while from January to April it is used for agricultural activities (Ahmad, Hossain, Pavelsky, et al., 2020; Rahaman et al., 2021). Conversely, the northwestern region, locally known as “Barind” is characterized by a semi-arid climate. Due to relatively less rainfall received over a higher mean elevation that allows little groundwater recharge, the Barind region of the northwest is often under frequent agricultural drought (Md. N. Hossain et al., 2016). The southeastern region is distinguished by its hilly terrain, a unique feature in the predominantly flat landscape of Bangladesh (S. Islam & Ma, 2018). The southwestern region is recognized as hosting one of the world's largest deltas, the Ganga-Brahmaputra-Meghna, and also the largest mangrove forest in the world – the Sundarbans (S. Islam & Gnauck, 2008).

To understand the surface water dynamics of each of these four regions, we identified the characteristics of its hydro-physical regime. The primary source of water influx to the lakes and wetlands in Bangladesh is precipitation that occurs during the distinctly wet season called the Monsoon comprising June to October. Thus, one of the key factors used to characterize the hydro-physical regime is the 30-year mean precipitation recorded during the month with the highest expected rainfall. Other contributing factors in defining a regime include temperature, which is a strong proxy for evaporation, terrain roughness, and the number of water bodies, as estimated from satellite data, present in the region.

To derive 30-year monthly mean values, we used both precipitation and temperature data. The influence of topography was incorporated through a measure of terrain roughness using Digital Elevation Models (DEM). Several methods have been proposed to estimate terrain roughness, such as standard deviation of slope (Ascione et al., 2008), relative Topographic Position Index (TPI) (Ebert, 2015; Riley et al., 2017; Tagil and Jenness, 2008), and slope variability (Ruszkiczay-Rüdiger et al., 2009). Our choice of method was guided by simplicity and past use case in the field of hydrology. Hence, we adopted the relative topographic position index. Several studies, such as those by Karimzadeh et al., (2015); Leonard et al., (2012); Riley et al., (2017) have shown the utility of TPI for classification and identifying hydro-physical regimes. TPI, developed by Weiss, (2001), compares the elevation of a cell relative to the average elevation of the surrounding cells. TPI values are straightforward and powerful tools for classifying landforms into morphological classes (Tagil and Jenness, 2008). TPI is calculated by the following equation:

$$TPI = \frac{Smoothed\ DEM - Minimum\ DEM}{Maximum\ DEM - Minimum\ DEM}$$

Additionally, we determined the number of water bodies in a region using Sentinel 1 (Synthetic Aperture Radar - SAR) imagery. To distinguish lakes and wetlands from rivers, we employed the Connected Component Analysis of Haors (CoCoAH) methodology (Ahmad et al., 2020). This technique allows for the differentiation of lakes and wetlands from rivers based on the eccentricity, perimeter, and area of a water body. Figure 3.3 illustrates the distribution of water bodies across various regions of Bangladesh.

Table 3.2 shows the values of the various estimates from the hydro-physical regime. Here, it is important to note the basis of defining the four distinct hydro-physical regimes based on factors that control lake formation and storage patterns. Each of these four regions are known to have distinct hydro-climatology and terrain features, as elaborated earlier. Widely-used classification systems such as Koppen climate (Peel et al., 2007) are not able to provide explanations for the observed distribution and pattern of surface water bodies, hence there is merit in deriving hydro-physical regimes. However, from Table 3.2, we can observe how each region can be explained of its distribution of water bodies as a function of selected key factors. For example, a region with the highest monthly precipitation with favorable terrain to form and store surface water (as defined by low TPI) is home to the highest density of seasonal lakes and wetlands (i.e. Regime 1). Similarly, a region with lowest monthly precipitation and unfavorable terrain for lake formation (as defined by high TPI and mean elevation) has the lowest density of lakes and wetlands (i.e., Regime 2). A region with very high precipitation but with steep slopes and mountainous terrain (as defined by very high TPI) similarly has lower lake and wetland density as the precipitation-driven runoff drains quickly to rivers and eventually to the ocean (i.e., Regime 3).

Figure 3.4 shows a few regions around the world, mostly deltas like Bangladesh, where the methodology developed in this study based on Bangladesh could be potentially reproduced, given limited number of *a priori* gauges, as a starting point for gauge placement optimization. Most of these regions are fertile deltas representing agriculturally productive systems and high population densities where surface water management is essential. This means that the tracking of surface water resources using citizen science and remote sensing is not only important but needs to be explored given the dense populations. It is important to stress that the global map in Figure 3.4 does not imply scalability of our proposed approach 'as is' but rather potential regions where such a methodology could be a potential candidate for exploring gauge network design for surface water tracking by building synergy with satellite remote sensing data from SWOT and nadir radar altimeter missions.

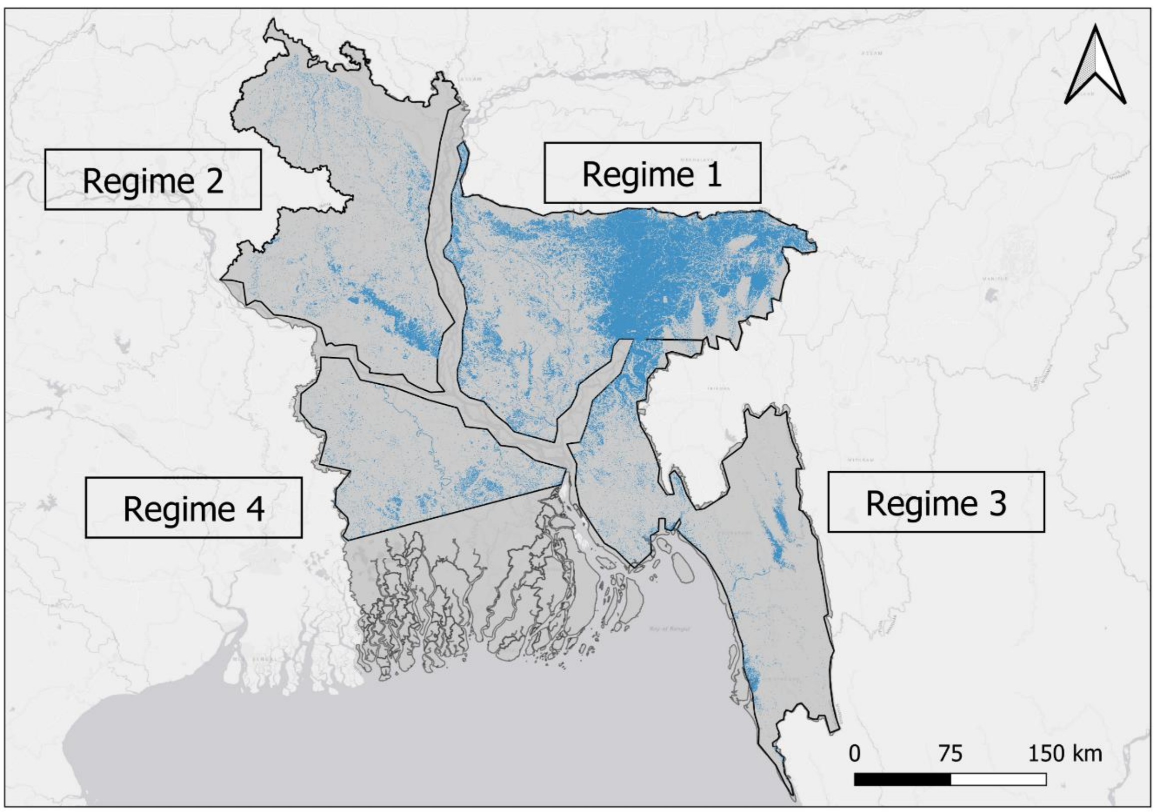

Figure 3.3: Identified water bodies in different Hydro-Physical Regimes defined by climatology in precipitation, temperature, TPI and water body density

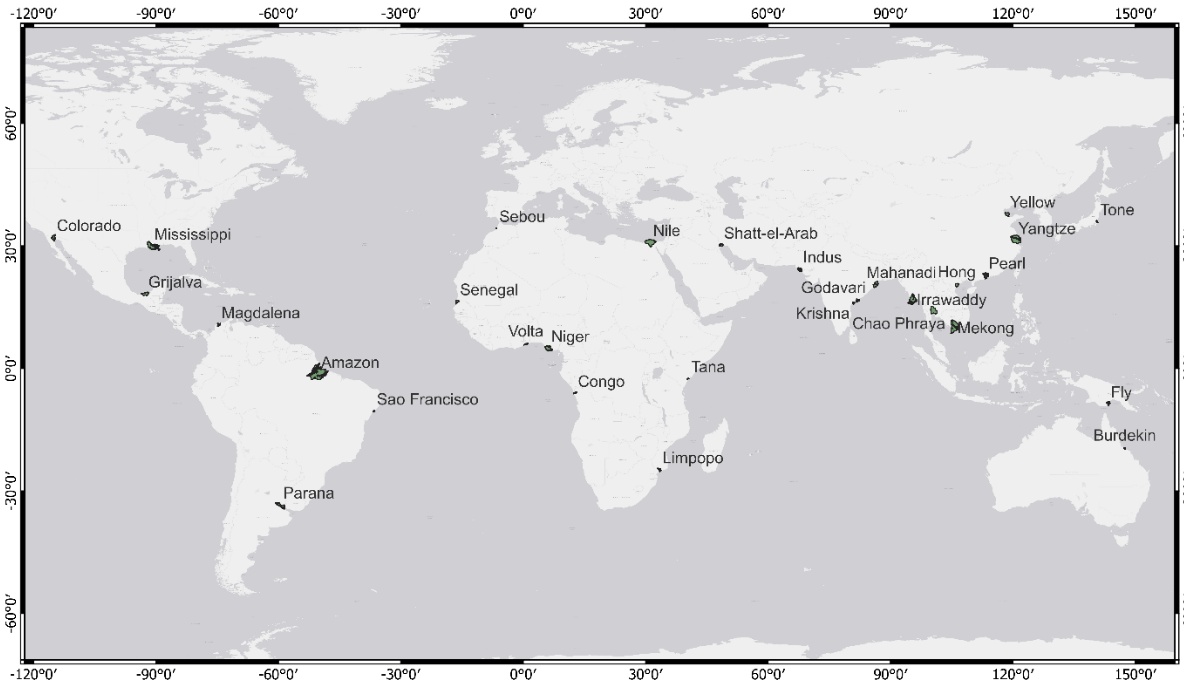

Figure 3.4: Map of popular deltas around the globe as the potential application of the proposed framework

Table 3.2: Hydro-Physical Regimes in Bangladesh

Hydro-Physical Regimes	**Precipitation (July)(mm)**	**Temperature (July) (°C)**	**TPI**	**Number of Waterbodies per 100 km^2**
Regime I Haor (1)	450	32.0	0.32	9
Regime II Barind (2)	350	30.0	0.39	1
Regime III Hilly Forested (3)	500	28.5	0.40	2
Regime IV Southwest (4)	340	30.0	0.33	3

3.5.2. A-V Curve Fitting and Monte Carlo Simulations

3.5.2.1. Volume Estimation of Waterbodies (Step 2)

The method to estimate the volume of the water bodies requires two input variables: water surface area and water surface elevation (Ahmad et al., 2020). The water surface area was determined through the integration of data from three remote sensing products, backscatter signal from Sentinel 1 and surface reflectance from Sentinel 2, and Landsat 8. The selection of these sensors was governed by their capability to estimate water surface area under various atmospheric conditions. An ensemble of water classification techniques was employed to refine the estimation of the water surface area. The ensemble was created utilizing Dynamic Surface Water Extent (DSWE; Jones, 2019) for Landsat 8 and Sentinel 2, Modified Normalized Difference Water Index (MNDWI; Xu, 2006) for Landsat 8 and Sentinel 2, and Sentinel 1 Backscattering. Thresholds selected for different water surface area estimation techniques are mentioned in Table 3.3: Water Surface Area Estimation Techniques This ensemble data, in conjunction with water surface elevation data submitted by citizen scientists, facilitated the estimation of stored water volume. Our methodology followed Khan et al., (2023), who demonstrated that the ensemble technique in volume estimation closely aligns with volume derived from high-resolution Planet 3m data, thereby establishing a benchmark for volume estimations. For an in-depth exploration of volume estimation, we direct the reader to Khan et al., (2023).

Table 3.3: Water Surface Area Estimation Techniques

Sensor	Revisit Time	Technique	Band (wavelength, micrometers)	Threshold
L8	16 days	Dynamic Surface Water Extent (DSWE)	Blue (0.45-0.51); Green (0.53-0.59); Red (0.64-0.67); NIR (0.85-0.88); SWIR1 (1.57-1.65); SWIR2 (2.11-2.29)	N/A
		Modified Normalized Difference Water Index (MNDWI)	Band 3 - Green band (0.53-0.59); Band 6 - Short-Wave Infrared (1.57-1.65)	0.3
S1	10 days	Backscattering thresholding	–	< -13 dB

S2	6 days	Dynamic Surface Water Extent (DSWE)	S2A - Blue (0.496); Green (0.56); Red (0.664); NIR (0.835); SWIR1 (1.613); SWIR2 (2.202) S2B - Blue (0.492); Green (0.559); Red (0.665); NIR (0.833); SWIR1 (1.610); SWIR2 (2.185)	N/A

3.5.2.2. Area – Volume (A-V) Curve Fitting (Step 3)

In order to generate an equation capable of estimating the volume of stored water in an ungauged water body, a regression line must be fitted between the water surface area and the volume of stored water. We propose two novel methodologies for hydrologic curve fitting: the Robust Curve Approach and the Hydrologic Parameter Derived Approach.

3.5.2.2.1. Robust Curve Approach

The underpinning assumption of the Robust Curve Approach is that an increase in the water surface area of a water body will correspondingly result in an increase in its stored volume of water. In this approach, area-volume data is sorted based on the area, selecting only those data points where the volume exceeds the preceding volume. Post-filtering, a linear regression is performed to fit the curve, omitting any constant, as the absence of detected water surface area logically corresponds to zero volume. Figure 3.5 illustrates the Robust Curve points and the linear regression fitted curve for regime 1 as an example. The logarithmic scale in Figure 3.5**Error! Reference source not found.**, offers a distinct perspective. Smaller water b odies, which were previously indistinguishable on the linear scale, now emerge with improved visibility. This transformation makes the range of smaller water bodies visually manageable, thus enhancing their comparability and decision-making during gauge network design.

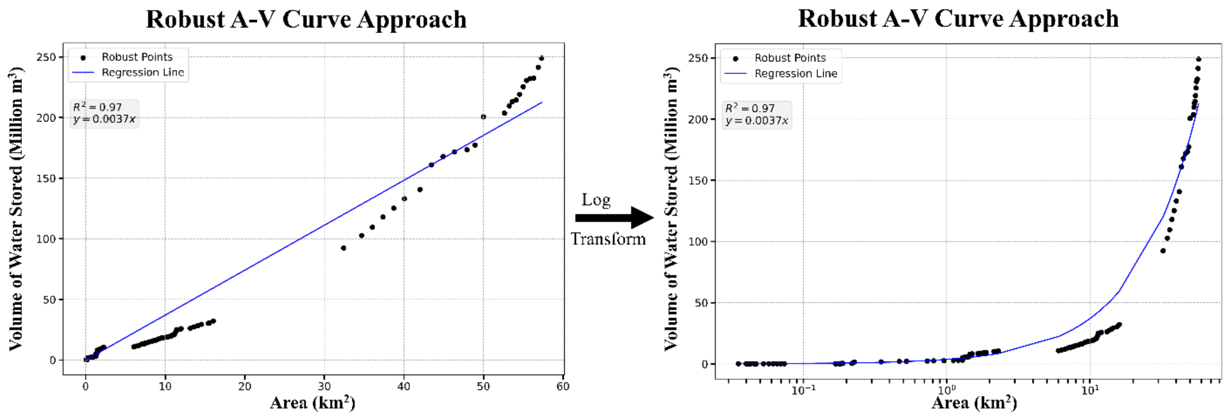

Figure 3.5: Robust Curve Approach. The figure shows the fitted Robust Curve with high goodness of fit ($R^2 = 0.97$) and the equation of line after log transformation of the x axis for the Hydro-physical regime I (Northeastern Bangladesh). The left panel depicts the curve in linear scale, whereas the right panel depicts the curve in log scale. The figure shows that the robust curve approach is a good predictor of volume for a given area when applied with log-transformed surface area of the water bodies.

3.5.2.2.2. Hydrologic Parameter Derived Approach (HyPD)

In the Hydrologic Parameter Derived (HyPD) approach we operate under the assumption that the volume of water stored in a water body is not solely dependent on the water body's surface area but is also influenced by an array of interconnected geophysical and climatic parameters. Our assumption acknowledges the complexity of hydrological systems where precipitation, temperature (proxy of evaporation), and the geophysical framework of the terrain collectively govern volumetric variations. Precipitation acts as the primary natural input, contributing towards the volume of lakes and wetlands. Conversely, evaporation represents a continual process of moisture loss, diminishing the water volume. The terrain, through its topographical and geological features, plays a critical role in containing the body of water, thereby influencing its volume. This recognition of multiple interdependent factors is critical in advancing our understanding and prediction of waterbody storage dynamics, ensuring that our estimations are robust and reflective of the underlying environmental processes. A multilinear regression was conducted to derive the area-volume curve relationship. Figure 3.6 depicts the area-volume data points and the fitted curve resulting from the multilinear regression. All the coefficients in the regression model yielded p-values less than 0.05, thereby affirming their statistical significance.

Further inspection of the HyPD curve fitting, as depicted in both linear and logarithmic scales in Figure 3.6, reveals insightful patterns. Upon examination of the logarithmic scale, it is apparent that HyPD exhibits curve fitting skill for smaller water bodies, but this skill decreases for larger bodies of water. This pattern is reinforced when examined on a linear scale where it can be observed that the HyPD curve tends to underfit the data after a surface area of around 20 to 30 km². The most probable explanation for this pattern can be attributed to the impact of terrain roughness on the volume storage of water bodies of varying sizes. It may be inferred that for larger water bodies, terrain roughness or climatology of precipitation and temperature does not significantly contribute to volume storage as much as it does for smaller water bodies.

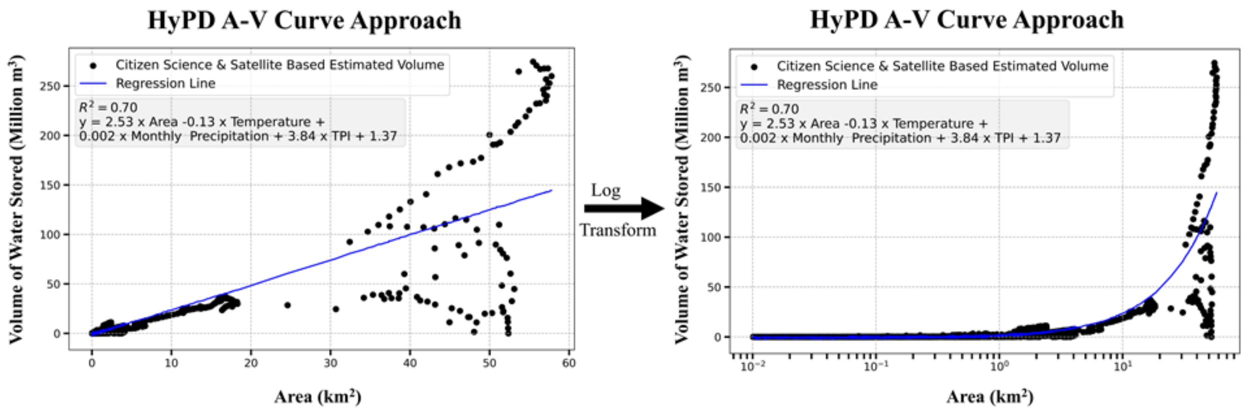

Figure 3.6: HyPD AV Curve Approach. The figure shows the fitted HyPD AV Curve with the goodness of fit ($R^2 = 0.70$) and the equation of line. The left panel depicts the curve in linear scale, whereas the right panel depicts the curve in log scale. As an example, this is shown for Hydro-physical Regime I (Northeastern Bangladesh)

In addressing the dynamics of area-volume (A-V) relationships in lakes and wetlands, our study employed Linear Regression and Multilinear Regression models. This choice was guided by the observed linearity in certain A-V relationships and the need to incorporate multiple influencing factors. While alternative models such as Ridge and Lasso Regression, Support Vector Regression (SVR), and Logistic Regression were considered, they were deemed less suitable due to our study's specific requirements for simplicity, interpretability, alignment with the nature of our data and global scalability. Linear and Multilinear Regression models provided the necessary balance between statistical robustness and clarity, aligning with the principles of hydrological modeling and the objectives of our study.

3.5.2.2.3. Hybrid Area – Volume Curve Approach

Careful visual inspection revealed that the HyPD approach exhibits a tendency towards marginal underestimation of the volume of water bodies associated with larger areas. Alternatively, the robust approach, as previously discussed, integrates data points that exceed the previously recorded volume data point in correspondence with the growth of area. Consequently, it takes into account the high area-volume points that are overlooked or under-represented by the HyPD approach.

In response to these findings, we propose the implementation of a Hybrid Area – Volume Curve Approach. This combines the beneficial aspects of both the Robust and the

HyPD approaches, thereby yielding a more precise and comprehensive evaluation of the area-volume relationship. The HyPD curve is better for smaller lakes and wetlands as compared to larger ones, while the robust curve fitting is able to fit larger waterbodies. The hybrid approach has been designed to effectively compensate for the inherent limitations of the individual methods. However, readers are advised to use their own judgement in determining the most suitable curve fitting method such as Robust Curve fitting, HyPD Curve fitting, Hybrid Curve fitting for their use case in their region. Figure 3.7 presents the Hybrid Curve fitting for Regime 1. While considering the Hybrid Curve, one should have a large number of dataset to apply the multilinear regression so as to implement the HyPD approach. If data scarcity is the problem, we recommend using Robust Curve only.

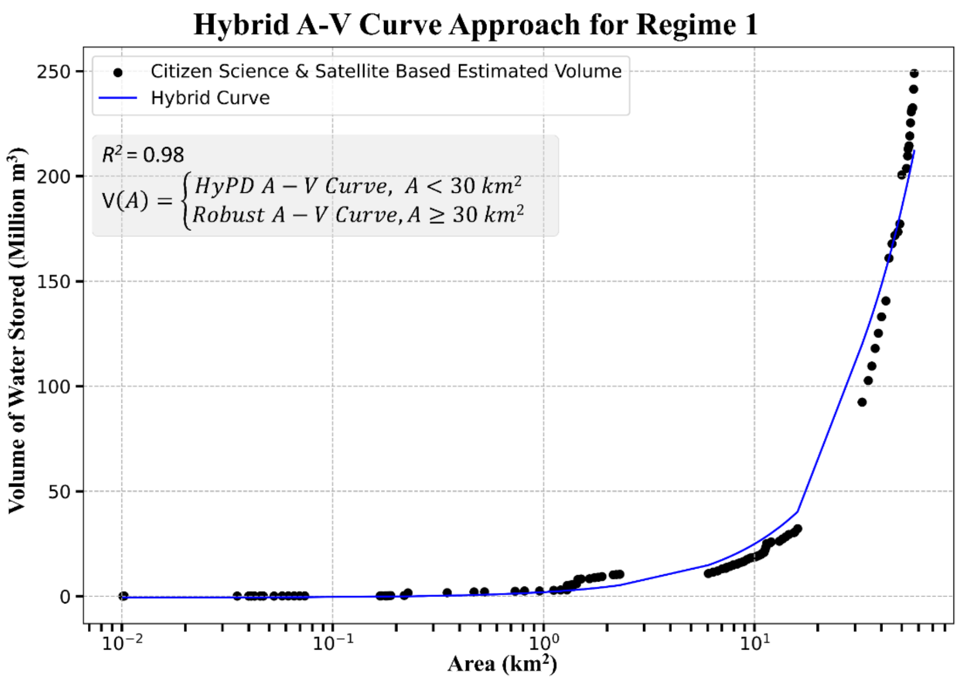

Figure 3.7: Hybrid Curve Approach. The Hybrid Curve Approach is the hybrid of Robust and HyPD curve. The figure shows the fitted line with a high goodness of fit ($R^2 = 0.98$) as compared to Robust and HypD curve fitting, As an example, this is shown for Hydro-physical Regime I (Northeastern Bangladesh)

3.5.3. Area and Volume Estimation of all the water bodies (Step 3)

We quantify the number of water bodies within the four hydro-physical regimes using the CoCoAH approach. Additionally, the water surface area pertaining to each of these water bodies was estimated, providing a comprehensive overview of water surface area within each regime. We estimate the volume of the water bodies quantified with the CoCoAH approach and the estimation of the water surface area described in 3.5.1. To estimate the volume of water stored within these bodies, we employed the A-V relationship derived from a hybrid of two A-V curve fitting methodologies. This approach facilitated the calculation of the cumulative volume of water stored in each hydro-physical regime, which served as a benchmark volume for each region.

3.5.4. Monte Carlo Simulations (Steps 4 and 5)

To perform sensitivity to gauge density (i.e, number of gauges) to volume change estimation, relative to the maximum number of gauges available, we implemented Monte Carlo simulations. This approach was premised on the concept introduced by Krajewski et al. (1991), who proposed to utilize a high-density precipitation gauge network as a benchmark, thereby enabling the assessment of other gauges relative to the overall network. In an environment containing N gauges, the selection of one, two and three random gauge generates ${}^{N}_{1}C$, ${}^{N}_{2}C$ and ${}^{N}_{3}C$ combinations, respectively. Figure 3.8 shows as an example the typical selection of gauges from the entire region of Bangladesh. However, in the methodology the gauges selection was implemented on each hydro-physical regime.

Monte Carlo simulation was performed to estimate the sensitivity of the number of gauges. Initially, a single gauge was selected from the N gauges at random, after which curve fitting was applied to all data points observed by this gauge, thereby enabling the estimation of the volume of water stored. This procedure was iteratively expanded by selecting two gauges at random from the N gauges and fitting a curve, as well as estimating the volume of water stored. The process was continued, each time adding one more gauge to the selection and performing curve fitting on all observed data points corresponding to the selected gauges. In each instance, the area-volume relationship was used to calculate the total volume of water stored, until the maximum density of gauges was reached.

In this study, the maximum number of gauges was 30, located in Regime I (northeast region). This resulted in ${}^{30}_{15}C$ (155117520) maximum number of combinations of gauges. Consequently, the total number of simulations was computed as follows:

$$Total\ Number\ of\ Simulations = {}^{30}_{1}C + {}^{30}_{2}C + {}^{30}_{3}C + \cdots + {}^{30}_{30}C = 2^{30} = 1073741824$$

Running 1,073,741,824 (~ 1 billion) simulations, fitting a curve for each simulation, and subsequently estimating the cumulative volume is a computationally intensive task. To ensure scalability and facilitate global application of our approach without necessitating high-performance computing machines. We limited the random selection of each gauge at each step to 20,000 times. This yielded a total of 600000 simulations. A comparison between the maximum possible number of simulations and the 600000 simulations revealed no substantial difference in the results.

Upon completion of each Monte Carlo realization, the percentage error in volume estimation was calculated using Equation (3.1).

$$Percentage\ Error = \frac{Benchmark\ Volume\ - Gauge\ Combination\ Mean\ Volume(Degraded\ Gauge\ Density)}{Benchmark\ Volume} \times 100 \quad (3.1)$$

The percentage error metric provided the flexibility to ascertain the minimum number of gauges that yielded an acceptable error in the region, with the threshold for acceptability defined by the user. Theoretically, with an increase in gauge density, the percentage error in cumulative volume estimation should decrease. Here, it is up to the user to define what is an acceptable level of percentage error in network design optimization.

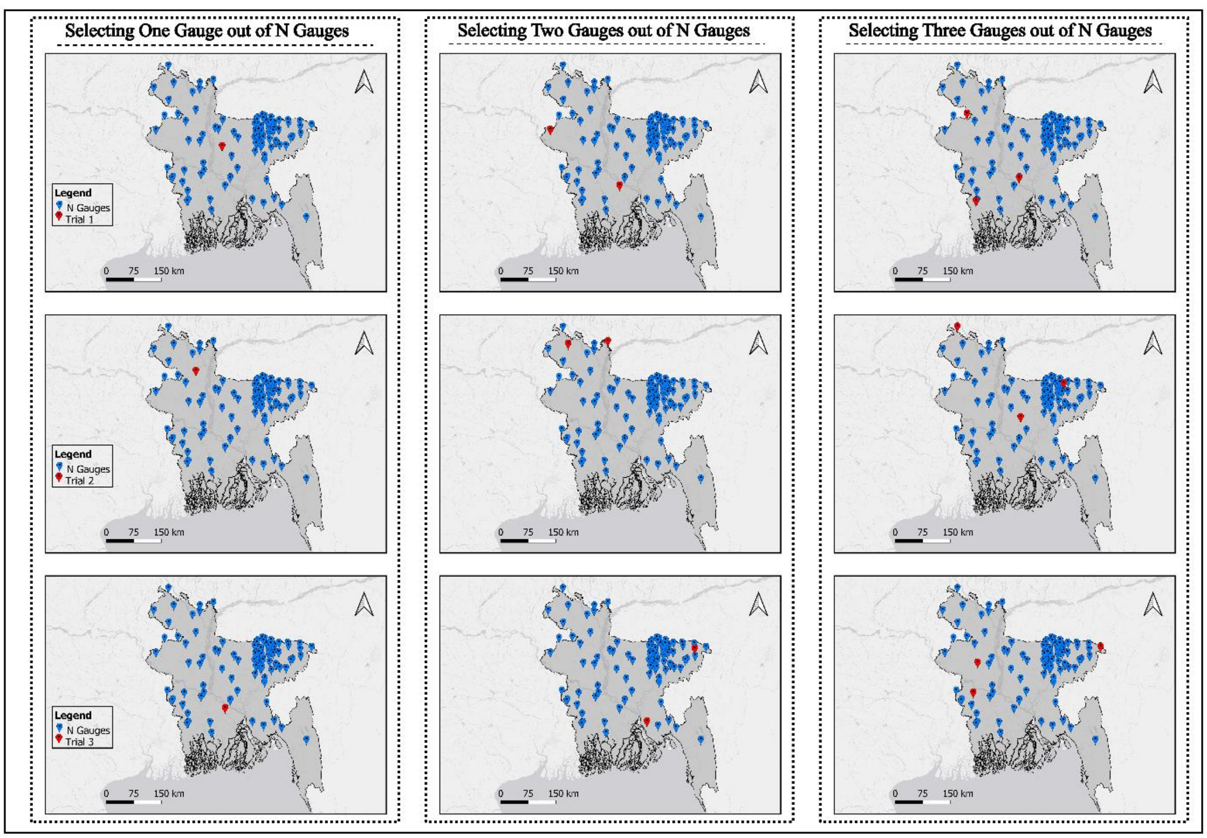

Figure 3.8: Pictorial representation of randomized selection of gauges for Monte Carlo simulations. The left panel shows the three combinations of randomly selecting one gauge out N gauges. The middle panel shows the three combinations of randomly selecting

two gauge out N gauges. The right panel shows the three combinations of randomly selecting three gauge out N gauges. The figure demonstrates scalability in two dimensions: extending horizontally (to the right) it represents the progression towards selecting all N gauges out of N, and vertically (downwards), it indicates the increasing number of trials or possible combinations of gauge selections.

3.5.5. Strategic Placement of Gauges (Step 6)

After identifying the optimal number of gauges, the subsequent challenge lies in their strategic or optimal placement as part of further optimization of the network. In other words, gauge network design is not just about number of gauges, but also about location of each gauge relative to the entire network. This placement must ensure the gauges' capability to adequately represent the given area. The selection of locations for gauges hinges on several factors, including the regional precipitation events, mean temperature, terrain roughness, and the gauges' contribution to the cumulative volume in other gauged locations sharing similar hydro-physical characteristics. To examine the individual contribution of each gauge towards the cumulative volume variations, we utilized the percentage error metric in conjunction with the gauge density plot. The primary goal of this approach is to determine which gauges are most vital for inclusion in the network, serving as a proxy for identifying locations driven by significant hydro-physical characteristics. The underlying rationale is that gauges recurring more frequently in these high performing subsets likely possess significant hydro-physical characteristics. Consequently, regions exhibiting similar hydro-physical traits should be prioritized during gauge placement.

3.5.6. Random Forest Classification and Probabilistic Plots

To identify potential gauge installation sites with combinations of temperature, precipitation, and terrain roughness conditions, we used random forest classification (Tyralis et al., 2019; Breiman et al., 1984). The random forest classification was tested and trained on two kinds of dataset. The first dataset addressed the hydro-physical characteristics of those gauges which have yielded high efficiency towards volume storage estimation. The second dataset addressed the hydro-physical characteristics of those gauges which had not yielded high efficiency towards volume storage estimation. Here the word ‘efficiency’ refers to the performance of the combination of gauges in capturing the ‘true’ total storage as estimated by the maximum set of gauges that were available (see Section 5 for more details). Following the creation of the random forest classifier model, raster data of the entire regime’s precipitation, temperature, and terrain roughness were used as input. This allowed us to estimate the likelihood of each pixel being preferred for gauge installation. This integrative approach of strategic placement allowed us to optimize the placement of gauges for a given regime by giving the likelihood of each site's effectiveness.

Employing the *RandomForestClassifier* from the scikit-learn library, we partitioned our dataset into a training set and a testing set at a 70:30 ratio. To mitigate the risk of overfitting, the maximum depth of the trees was capped at two levels. This model was independently developed and validated for each regime.

The accuracy score and F1 scores for these models are presented in Table 3.4. Caution is warranted in interpreting the absolute numeric values of the model results because no model is perfect, which Table 3.4 implies for Regimes I and II. Given the relatively small sample size of gauges (less than 100) and despite limiting the depth of trees, there is likely a high likelihood of overfitting. It is well known that in random forest classifier models, decision trees run the risk of overfitting as they tend to tightly fit all the samples within training data. In our case, the 70:30 ratio for training and testing results in very limited number of data points, making the model prone to overfitting. Nevertheless, in relative terms, the results inform us that models for Regime I, II, and IV exhibit high accuracy and F1 scores, indicating relatively more robust classification performance when there are simply more gauges. Conversely, the model pertaining to Regime III demonstrated a lower performance metric, which can be attributed to the limited number of gauged water bodies available in the particular regime.

Table 3.4: Random Forest Model Accuracy and F1 Score for Each Regime

Regime	**Model Accuracy**	**F1 Score**
Regime I	1	1
Regime II	0.75	0.86
Regime III	0.5	0.67
Regime IV	1	1

3.6. Results and Discussion

3.6.1. Estimating benchmark volume for each hydro-physical regime

Figure 3.9 provides the amount of surface water stored in different regimes in a logarithmic scale to permit a holistic comparison across all the regimes. The ensemble technique has been used to estimate the mean volume of surface water, represented by the shaded bar in the figure. The same ensemble technique helped us estimate the maximum and minimum amounts of surface water volume, represented by the uncertainty bars. For more information about the ensemble method and how we dealt with the uncertainty, please refer to the study by Khan et al., (2023). To estimate the benchmark volume, the analysis was performed for the month of July, the time of the year when Bangladesh is mostly inundated with water. We found the larger volume of water is typically stored (31 km^3) in Regime I. This regime has gathered considerable attention from researchers due to its susceptibility to severe monsoon-induced flooding and its high agricultural productivity (Haque et al., 2021; Islam et al., 2021; Kamruzzaman and Shaw, 2018). This result is consistent with the work from Ahmad et al., (2020), who using a combination of gauges and virtual gauges from satellite altimeters, projected that the total volume of water stored in Northeastern Bangladesh could reach up to 30.8 km^3. This convergence of estimates using a citizen-science framework lends further credibility to our findings, affirming that our volumetric assessment falls within an acceptable range. However, other regimes present a dearth of estimates for the total volume of surface water stored, thereby precluding us from contrasting our volumetric estimates for these regimes.

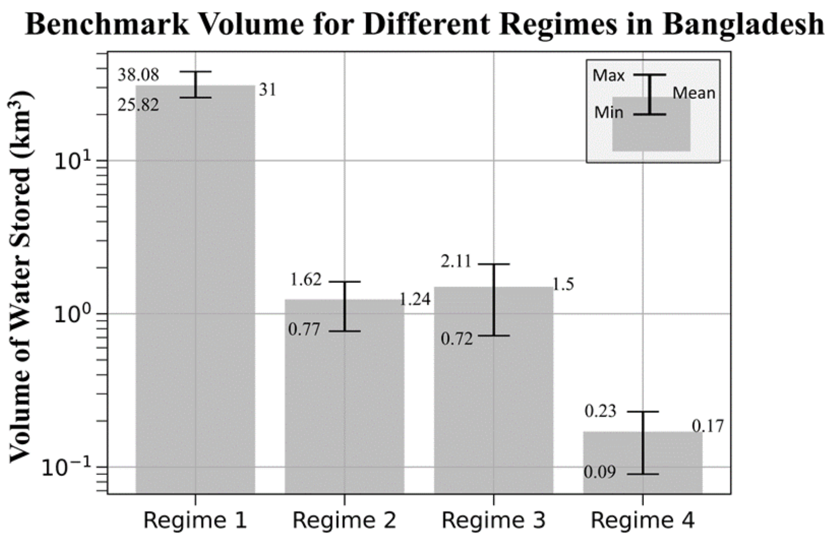

Figure 3.9: Benchmark volume of water stored in different hydro-physical regimes of Bangladesh. See Table 3.2 for the classification of regimes.

3.6.2. Optimum number of gauges

The execution of Monte Carlo simulations facilitated the determination of the sensitivity of number of gauges to water volume estimations relative to what is possible to decipher based on the maximum number of gauges available in this study. The results shown in Figure 3.10 and Figure 3.11, represent the percentage error of gauge combinations derived from Monte Carlo realizations. The shaded area in the plot represents the 90-percentile range, spanning the 5% to 95% quantile values. The mean of the gauge combinations for specific number of gauges is illustrated by the dashed line.

Figure 3.10 and Figure 3.11 provide stakeholders with the capability to define an acceptable margin of error for their volumetric estimations. Gauge combinations with low percentage error in Figure 3.10 and Figure 3.11 are the gauges that can estimate the cumulative storage efficiently and vice-versa. High-efficiency gauges are the gauges that have a percentage error less than 10% (for regime I,II,III) and a percentage error of less than 20% for regime IV (due to the slope of the Regime IV of Figure 3.11). These figures can then guide the determination of the optimal number of gauges to install given that there is likely to be limited number of gauges available that will be less dense than what was available in this study. This approach empowers water managers to incorporate both citizen science and satellite-derived data in their decision-making processes, allowing them to understand how gauge density impacts the estimation of volume change, corresponding to their tolerable error range, which is flexible in our proposed methodology.

In our subsequent analysis, including sensitivity analysis and threshold selection, we used a 10% error as an example of acceptable margin for all the regimes, though a user could certainly choose other thresholds. Therefore, for Regime 1 (Haor), which is shown in the top panel of the figure and has a total land surface area of about 33367 km², requires 28 gauges to represent the area with 90% accuracy. On the other hand, Regime II (northwest region of Barind), with a total land surface area of 30411 km², necessitates 11 gauges for a 90% accurate representation. It is important to remember that all this is relative to the maximum gauge density that was available in this study which in itself is incapable of yielding the true volume change estimations as increasing the density will likely keep improving the volume estimations until every single lake is gauged. Similarly, from the Figure 3.11, we deduced that 5 of 6 and 9 of 9 gauges are quite sufficient to accurately represent Regime III (Hilly and Forested) and Regime IV (Southwest), with land surface areas of approximately 27045 km² and 15648 km², respectively. Notably, Regime IV demands the installation of all gauges to achieve a minimum accuracy of 90%, possibly due to the high variability of the climate near coastal regions, or it might indicate that Regime IV requires additional gauge installations for a more accurate representation.

The findings from Figure 3.10 and Figure 3.11 have one caveat. None of the regions indicate a complete plateauing effect, when zoomed in, as the number of gauges reaches the maximum, thus suggesting that if more gauges were available in each region, the percent error would likely continue to decline. The likely ‘ideal’ maximum number of gauges after which the estimation of surface water storage variations show insignificant improvement is practically impossible to know without actually placing numerous gauges to intensify the benchmark network. This is an inherent limitation of any gauge network design study such as the one presented here and one should focus more on the sensitivity of number of gauges to volume estimation for each region.

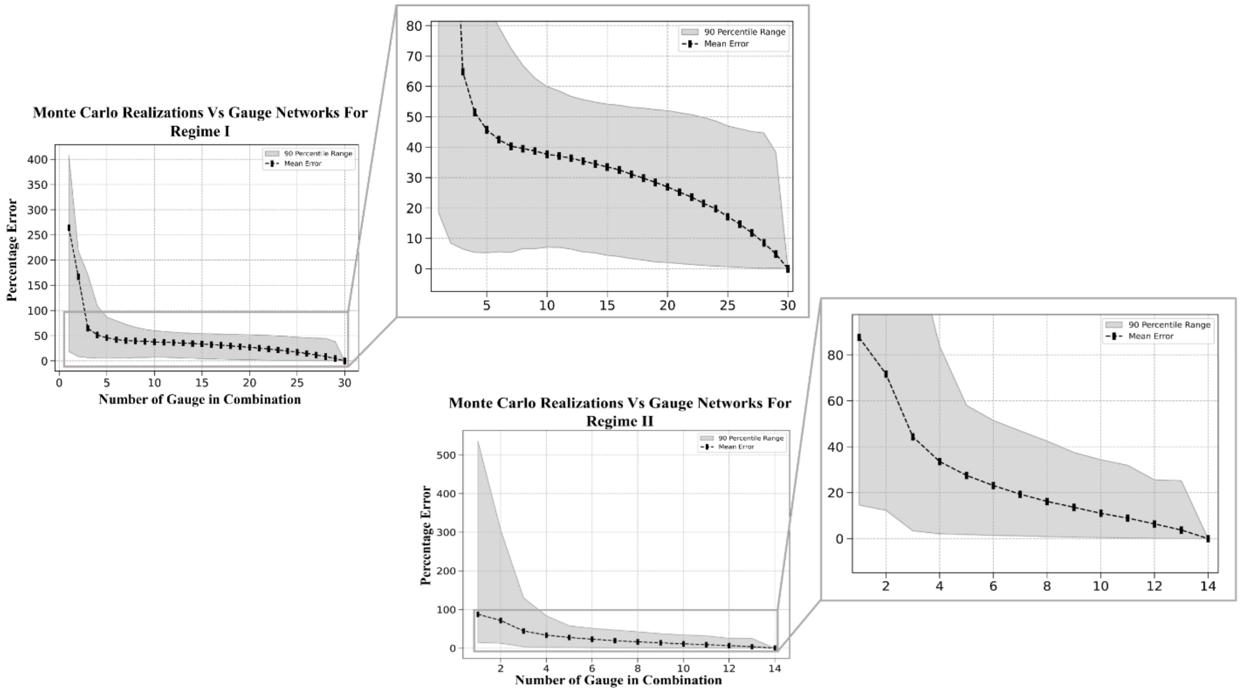

Figure 3.10: Monte Carlo Realization Vs Gauge Network for Hydro-physical Regimes I and II. Each panel represents the percentage error of gauge combinations derived from Monte Carlo realizations. The shaded area in the plot represents the 90-percentile range, spanning the 5% to 95% quantile values. The mean of the gauge combinations for specific number of gauges is illustrated by the dashed line.

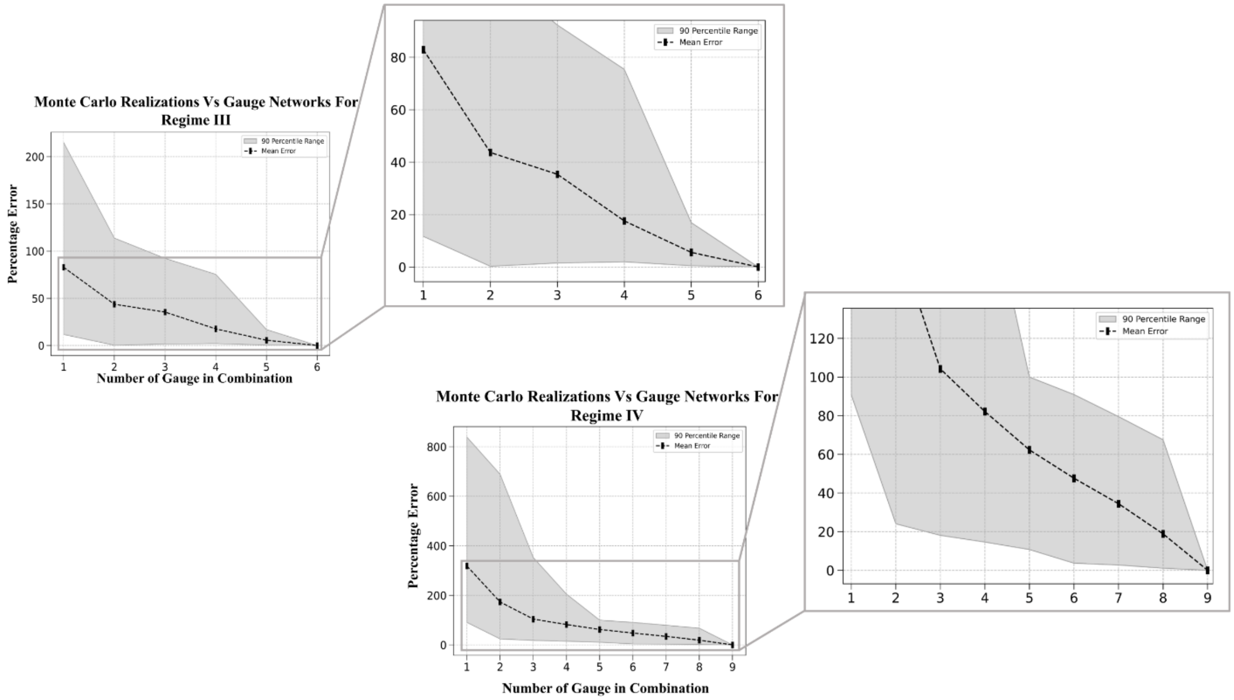

Figure 3.11: Monte Carlo Realization Vs Gauge Network for Hydro-physical Regimes III and IV

By carefully studying the Figure 3.10 and Figure 3.11 we can summarize following insights.

1. Decrease in Error with More Gauges: Figure 3.10 and Figure 3.11 showed a clear trend of decreasing error as the number of gauges increases. This suggests that using a greater number of gauges in combination leads to more accurate estimations of the cumulative volume of water.

2. Non-zero Derivative (Slope): From a calculus perspective, Figure 3.10 and Figure 3.11 does not reach zero slope, which would indicate no change in error with respect to additional gauges. The continued decline of the error suggests that the derivative remains negative. The lack of a zero slope suggests that extrapolating beyond the current maximum number of gauges may continue to show a reduction in error. If more gauges were to be considered, it is reasonable, based on the current trend, to expect continued, albeit small, improvements in estimation accuracy showing diminishing returns.

3. Practical Implications: The observed trend indicates that there is still potential value in adding more gauges for more precise volume storage estimations. The decision-makers would have to consider the cost-benefit analysis, taking into account the diminishing rate of improvement, to decide on further expansion of the gauge network.

In examining the shape, and slope of the relationship between the number of gauges and the percentage error in cumulative volume estimations, the data reveals a consistent reduction in error with increasing gauge density. This trend persists across all regimes, with no indication of a plateau within the studied range. Given that not all water bodies have been gauged, the potential for further reduction in error exists. However, this potential must be balanced against resource constraints, prompting an optimization challenge. As gauge numbers rise, the slope's trajectory suggests that while benefits continue, they do so at a decreasing rate. Therefore, strategic decisions must consider the trade-off between the marginal accuracy gains and the costs associated with gauge deployment.

3.6.3. Strategic Placement of Gauges

We used a random forest classification to generate a probabilistic plot, as depicted in Figure 3.12 that maps Bangladesh and indicates the potential impact of gauge installation at various locations. This plot reflects the likelihood, denoted as 'P', of a gauge installed at a given location - characterized by specific precipitation, temperature, and terrain roughness parameters - contributing significantly to the estimation of the cumulative volume of surface water stored. For all the regimes we considered 90% accuracy, however for Regime IV we considered 80% accuracy. This distinction was important as including all gauges would have led the random forest algorithm to incorrectly classify all locations as favorable (discussed in detail next).

In the northeastern region of Bangladesh (Regime I - Haor), the Figure 3.12 reveals the presence of certain areas where gauge installation could markedly enhance our cumulative volume estimation. This region, given its high density of water bodies, presents ample opportunity for impactful gauge placement, or the highest impact in capturing water storage variability. Conversely, the northwest (Regime II - Barind), a region with fewer water bodies, also indicates high contribution from most of the potential gauge installations. This region had less gauges to begin with and our network design methodology is able to indicate that placing more at many strategic locations could dramatically improve surface water volume estimation. The southeastern region of the Figure 3.12 (Regime III), which is characterized by hilly terrain, records the highest precipitation levels in the country. Installation of gauges in this region could yield valuable data and a greater return in investment. The southwestern region (Regime IV) reveals a heterogeneous distribution of probabilities for gauge placement. Some areas exhibit higher probabilities than others, reflecting the varying potential for impactful gauge installation. Due to this high variability we believe that the Regime IV considered all the existing gauges as optimal gauges. Overall, the key summary that emerges from the use of our proposed methodology to identify strategic locations of gauge placement is as follows. Wherever there are large regions with a very high probability of occurrence (close to 1), placing additional gauges in any combination there would be a good investment for a citizen science-satellite framework for surface water storage monitoring. However, it is crucial to note that gauge placement decisions are multifactorial. Apart from hydro-physical factors, socio-economic variables such as population density and feasibility of gauge installation also play pivotal roles. Thus, an integrated approach that takes into account both these aspects is essential for effective gauge placement.

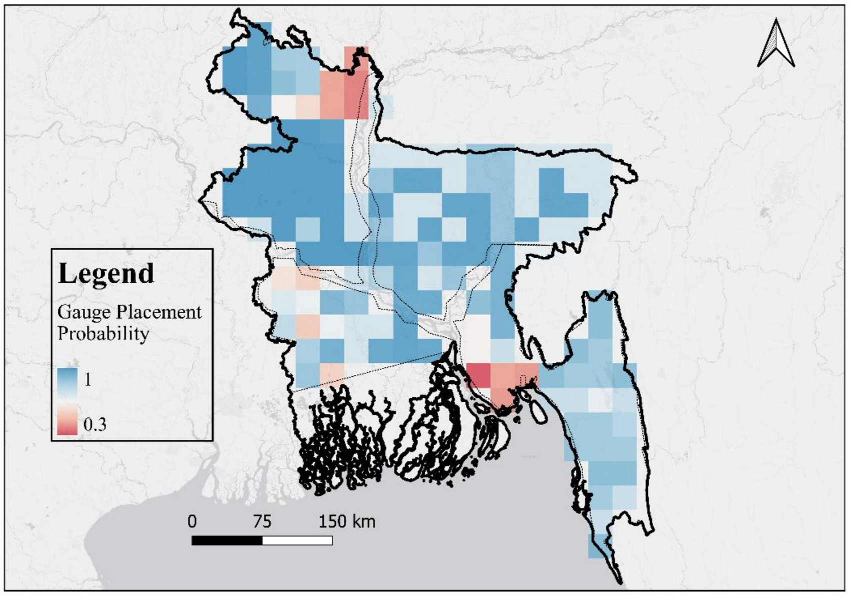

Figure 3.12: Probabilistic Plot for Gauge Placement in Bangladesh. The highest preferred locations are numbered as one and are marked with 'blue' color while the less or not preferred locations are marked with 'red' color. With the increase in the value, the preference of a location over other locations increases. The map was generated based on a percent of error of 10% as an acceptable value.

3.7. Conclusion

This study proposed a systematic method for understanding the sensitivity of lake gauge density for optimal network design for Bangladesh, given limited number of a priori gauges, within a citizen science remote sensing framework such as LOCSS. Since it is impossible to gauge every single lake using on-ground sensors, such a study can enhance surface water management through optimal use of resources afforded with the synergistic use of satellite remote sensing, citizen scientists and on-ground sensors. Our study can be summarized in seven lessons for potential users wishing to implement our proposed methodology for network design in their region of interest:

Lesson 1	We designed a scalable, integrated optimum gauge network framework that can adequately represent a region, utilizing both citizen-science and satellite-derived datasets. The framework can accommodate any form of water elevation data, not limited to those

	recorded by citizen scientists. Bangladesh was selected as the test case for applying this framework due to high density of gauge networks.
Lesson 2	Temperature, precipitation, and terrain roughness, were incorporated to estimate the optimal number of gauges for adequate regional representation. The approach remains flexible, and can be extended to consider other factors like snowmelt, streamflow, population density, infrastructure development, and other physical and socio-economic determinants. Regimes were created using the hydro-physical factors, and the understanding of regime can be extended to new regions planning to install gauges.
Lesson 3	Based on the relation area-volume, we applied two approaches, the Robust Curve and Hydrologic Parameter Derived (HyPD) methods, for fitting the Area-Volume Curve. The hydrologic community can select either curve or both based on their judgement of the region to fit Area-Volume data points.
Lesson 4	The benchmark volume in each regime was estimated, which allows water managers worldwide to use this framework for understanding the total volume of surface water stored in their respective countries - a crucial aspect of water management and resource placement.
Lesson 5	The framework of optimum number of gauges does not provide a single solution but a range of optimal solutions, leaving the ultimate decision on network adequacy to water managers.
Lesson 6	The probabilistic plot identifies numerous locations where gauge installation would yield the maximum return on investment. Water managers can couple this plot with other factors such as population density maps to choose gauge installation sites that maximize return on investment and foster local understanding of lake and wetland

	behavior. This could initiate data collection through citizen science in the region.
Lesson 7	The techniques and results developed in this study, such as the area-volume curve fitting and benchmark volume estimation, can be potentially utilized to assess the SWOT lake products regionally. While SWOT provides water surface area and elevation data, integrating it with proposed curve fitting techniques and volume estimation using citizen science-based gauges can allow for higher frequency volume storage estimations for small lakes and wetlands.

Addressing the pressing issue of effective surface water management, this study presents a structured approach that countries, particularly those suffering from water scarcity or possessing ample water resources but lacking efficient management strategies, can adopt. Bangladesh, identified by the Intergovernmental Panel on Climate Change (IPCC) as climate vulnerable due to recurrent droughts and flooding, and having high gauge density of network of gauges from the LOCSS network, is used as a test case in this research. Effective management of even the smallest water resource units can make a significant difference in such countries, and the implementation of gauges is a key step in this process within the framework of citizen science and remote sensing missions such as SWOT. By leveraging the high lake gauge density and the uniquely diverse and dynamic lake/wetland behavior, we have used Bangladesh as a test-bed for developing a methodology for understanding the sensitivity to gauge density and its optimal placement for volume estimation given a priori gauges and the region-specific characteristics.

The findings of our study are not without limitations. Our entire study is conditioned on the premise that the current network of dense LOCSS gauges in Bangladesh offers a reasonable benchmark to compare against. Furthermore, we assume that a higher density of gauges is practically unfeasible in most places around the world, including the developed regions. This premise should be questioned and not overstated when possible where a more superior gauging density for lakes is available. Additionally, while temperature, precipitation, and terrain were significant determinants in our study, we acknowledge that they do not encompass the full spectrum of factors influencing water volume changes. Omitted aspects such as groundwater seepage and lake leakage, due to their minimal contribution within our study's scope, might demand consideration in areas where they exert a greater influence on hydrological assessments. The HyPD approach, while robust within certain datasets and environments, may not universally apply, especially in regions constrained by data paucity or where empirical performance metrics suggest otherwise. In such instances, the more foundational Robust A-V curve methodology, predicated on straightforward hydrological premises, could provide broader applicability. Nevertheless, we believe that our recommended approach for developing a reproducible methodology for lake gauge network design, relative to the Bangladesh benchmark, has merit in application around the world given the limited and often declining lake gauging and greater availability of satellite remote sensing data and citizen scientists.

The monitoring framework developed in Chapter 3, together with the methodology developed in Chapter 2 for estimating water volume in water bodies using satellite and citizen science data and quantifying the associated uncertainties, establishes a comprehensive approach for integrated surface water monitoring (monitoring of *source*, see Section 1.1.4 of Chapter 1). This framework integrates scientific analysis, technological tools, and citizen participation to support the sustainable management of surface water resources. In Chapter 4, the focus will now shift from the *source* to the *sink* - that is, the application of surface water resources, where the water stored in *sources* is utilized (hence called *sink*). Since irrigation accounts for the largest share of surface water consumption, optimizing irrigation canal networks offers the greatest potential for enhancing overall water sustainability. Efficient allocation of water in this sector can significantly improve resource use efficiency and contribute to potential food security. Chapter 4 presents a scalable framework for optimizing irrigation canal operations, providing guidance to canal managers for allocating water in accordance with net crop water requirements.

Chapter 4. Satellite Data Rendered Irrigation using Penman-Monteith and SEBAL - sDRIPS for Surface Water Irrigation Optimization

4.1. Abstract

This study proposes a satellite remote sensing-based water-provider-centric irrigation advisory system designed to manage surface water resources and allocate water efficiently to areas in need, thereby promoting sustainable irrigation practices in the context of a changing climate. The system utilizes satellite remote sensing based SEBAL (Surface Energy Balance Algorithm for Land) and Penman-Monteith evapotranspiration models to estimate crop water use. By integrating the responses from the previous irrigation cycle, current precipitation, forecasted precipitation, and evapotranspiration-based water needs, the framework calculates the net water requirements for command areas within irrigation canal networks. Operating on a weekly basis, the system generates advisories that enable the irrigation water provider to make informed, science-based decisions about water allocation. These advisories quantify the net water requirement, giving water providers the flexibility to dispatch water to areas of higher need based on their on-ground judgment. Additionally, the proposed framework can simulate future cropping patterns by assuming potential policy changes or net reduction in water supply in the main canal due to climate change or increased transboundary withdrawal. The advisory system is co-developed and implemented with the irrigation management agency called Bangladesh Water Development Board on the Teesta River Irrigation System located in Northern Bangladesh. The study demonstrates its effectiveness when compared against actual water supplied for irrigation. However, the application of sDRIPS is not limited to Bangladesh, as it is scalable to other regions with similar water management challenges for agriculture.

Note: This chapter has been submitted in its current form as an article in *Hydrology and Earth System Sciences (HESS).* The only differences are in the numbering of figures, tables, and references to other sections.

4.2. Introduction

Freshwater is an essential resource for human survival and underpins a wide range of economic activities. Among the various sectors that are dependent on freshwater, agriculture is the largest consumer, accounting for approximately 85% of all human freshwater water consumption (D'Odorico et al., 2020). Projections for the year 2050 indicate a global population increase to more than 9.0 billion, which will further exacerbate the pressure on existing freshwater resources for food production, particularly groundwater (Boretti & Rosa, 2019). Recent studies have highlighted the significant depletion of groundwater resources in major food-producing regions and other areas worldwide (Feng et al., 2013; Jasechko et al., 2024; Kuang et al., 2024; P.-W. Liu et al., 2022). In response to this issue, there is an urgent need for the conjunctive and optimal use of groundwater and surface water resources. Nations such as India, Pakistan, and Bangladesh have increasingly promoted the use of surface water resources to alleviate pressure on groundwater reserves (Paul & Hasan, 2021). To enhance the effectiveness of these efforts, systematic and frequent water accounting is essential. This process should comprehensively account for direct human water use—including agricultural, municipal, commercial, and industrial consumption, as well as indirect losses due to reservoir evaporation and evapotranspiration (Richter et al., 2024). Such analyses are critical for quantifying stress on different water sources and informing strategies for their sustainable and efficient management to meet growing water demands.

Assuming a scenario where both surface water and groundwater resources are sufficient to meet food production needs, and we have a robust infrastructure for supplying from both sources, surface water is the more sustainable source due to its multifaceted benefits. Firstly, surface water often contains higher nutrient levels than groundwater, which can lead to increased crop yields. Secondly, by reducing the reliance on groundwater pumping, surface water allows aquifers to recover naturally by acting as a "storage battery" during dry seasons and ensuring a long-term water supply for agricultural and other uses. Thirdly, and perhaps most importantly, surface water irrigation systems are mostly gravity-driven and require less energy than groundwater pumping systems. This reduction in energy consumption translates to lower greenhouse gas emissions, thereby reducing the carbon footprint of agricultural practices (Qin et al., 2024). Accounting for government policies, over-exploitation of existing irrigation water resources, and the benefits of surface water irrigation, there is a need to manage surface water resources sustainably and efficiently.

To manage surface water irrigation more efficiently, adapting traditional irrigation practices to address the challenges posed by climate variability and increasing demand is crucial. Rising temperatures and changing rainfall patterns significantly impact crop growth and surface water availability. This in turn, demands that current surface water irrigation practices become more optimized. Studies led by Kovenock and Farhat (2018, 2021) highlight the complex relationship between rising temperatures and biomass production. Kovenock &

Swann (2018) have reported that increased carbon dioxide (CO_2) concentrations can negatively affect biomass productivity and alter the leaf area, which reduces the resilience against climate change. Farhat et al. (2021) showed that the elevated temperature leads to a higher uptake of toxic substances in groundwater such as arsenic. This bioaccumulation eventually results in plant failure at elevated temperatures, affecting both food quality and quantity. In many places, such as in Bangladesh where groundwater can often be contaminated by arsenic (Pogorski and Berg, 2020), the surface water usually has a better quality for irrigation applications. Overall, surface water can present a better alternative, in many places, in the light of climate change, for mitigating these adverse effects on food production in terms of quality and sustainability.

To fully leverage the benefits of surface water and adapt to climate variability, it is essential to overcome the limitations of traditional irrigation practices. Traditional-irrigation-practices specifically refer to those employed by the water-providers, encompassing temporal, intensity-based, and spatial components of irrigation decision-making. These components are broken down into the following:

1. When to irrigate (time - T) – determining whether water should be distributed in the current cycle based on the combined effects of the previous irrigation cycle, precipitation response, and crop water need;
2. How much water is needed (intensity – L^3) – evaluating the volume of water required by analyzing the intensity of the combined effects of the previous irrigation cycle, precipitation events, and crop water need; and
3. Which region to irrigate (space - Location) – identifying regions requiring irrigation by assessing the spatial distribution of the previous irrigation cycle, precipitation events, and crop water need.

In reality, traditional irrigation practices are not limited to water-providers but extend to water-consumers (farmers), who rely heavily on traditional irrigation calendars. These calendars, which are static and developed based on the historical climatology of the previous century assume a fixed cropping pattern. These calendars inform farmers of the irrigation schedule to follow but are becoming increasingly inadequate in the face of rising food production demands. The reliance on these static crop calendars leads to inefficiencies in water management and reduced agricultural productivity, as evidenced in many countries such as India, Bangladesh, Pakistan, and Australia (Bose et al., 2021; Bretreger et al., 2020; F. Hossain et al., 2017). Regardless of the region, farmers typically are not adaptive to rapid climate variability and cannot accurately predict precipitation events and their amounts at sub-seasonal irrigation timescales. This uncertainty often leads to overwatering of the fields to avoid the risk of crop loss due to rising temperatures. Overwatering, however, can inadvertently wash away essential soil nutrients, while underwatering can stress crops, both resulting in reduced yields.

These practices lead to inefficiencies in water use and diminish soil fertility and agricultural productivity (Hossain et al., 2017).

Considering the effects on agricultural yield due to altered rainfall patterns and rising temperatures, there is an urgent need for a scientific-data-driven approach to optimize existing irrigation practices to achieve maximum crop water productivity. The potential of using satellite data for sustainable agriculture and efficient irrigation has been widely acknowledged in recent years (Deines et al., 2019; Sishodia et al., 2020; C. Zhang et al., 2022). Several studies have employed crop coefficients (Kc) based evapotranspiration techniques to enhance the understanding of irrigation schedules and irrigate fields based on actual water needs (Bretreger et al., 2020; Gabr & Fattouh, 2021). Tools like the Food and Agriculture Organization's (FAO) CROPWAT, which utilizes the Penman-Monteith equation (Allen et al., 1998), have provided valuable insights into regional water requirements (Dong, 2018; Khaydar et al., 2021; Solangi et al., 2022; Surendran et al., 2015).

Despite significant advancements, substantial gaps remain in the current body of research that need to be addressed. These gaps arise from limitations in the temporal and spatial resolution of past studies. The trade-off between temporal and spatial resolution of satellite data makes it challenging to estimate evapotranspiration (ET) at high temporal and spatial resolutions simultaneously. Typically, temporal resolution ranges from daily to monthly time steps. When the temporal resolution is one day, the spatial resolution is usually at the kilometer scale (Chen et al., 2021). Conversely, if a medium spatial resolution sensor having 30-meter pixels is selected, the temporal resolution is typically around 16 days from polar-orbiting optical satellites (Tasumi, 2019). To achieve an optimal balance between spatial and temporal resolution, a combination of satellites needs to be used. Additionally, challenges exist in operationalizing findings or tools due to the latency of the available latest meteorological data. For instance, the Global Land Data Assimilation System (GLDAS), which provides global meteorological forcing data, can be used in estimating evapotranspiration (Pareeth & Karimi, 2023), has an early product stream with a latency of about 1.5 months. This latency hinders the ability to use these datasets for near real-time operational studies.

Tools available for surface water irrigation also suffer from an inability to accurately distinguish between the actual water provided to plants and the water needed by plants. This arises from the difficulty in estimating ET under ambient water stress conditions for each crop and lack of metering and monitoring. The incomplete information makes it difficult to estimate net water requirements and project potential water savings. Additionally, there is a lack of integration of precipitation forecasts at the operational level to optimize irrigation even further. It is important to acknowledge that, to the best of our knowledge, there is no freely available open-source tool that addresses these gaps in optimizing surface water irrigation schemes based on satellite data.

In this study, we present Satellite Data Rendered Irrigation using Penman-Monteith and SEBAL (sDRIPS., pronounced as "drips"), a cloud-based optimized surface water irrigation advisory system that leverages earth observations and weather models. sDRIPS. utilizes publicly available data from the Landsat satellite series, the Global Forecast System (GFS), and the Global Precipitation Measurement's (GPM) IMERG data (precipitation dataset). The founding concept of sDRIPS originated from the Irrigation Advisory System (IAS) first prototyped by Hossain et al. (2017). Later, Bose et al. (2021) integrated Gravity Recovery and Climate Experiment (GRACE) satellite-based water storage data with Landsat data to quantify the impact of IAS. Bose et al. (2021) also used satellite and model weather data to estimate ET using the Penman-Monteith method (Allen et al., 1998) and the Surface Energy Balance Algorithm for Land (SEBAL; Bastiaanssen et al., 1998) as proxies for crop water demand and actual water consumed, respectively

The core concept of IAS has been implemented in various regions under different names, such as IAS for Pakistan (F. Hossain et al., 2017), Provision for Advisory on Necessary Irrigation (PANI) in India (F. Hossain et al., 2020), and Integrated Rice Advisory System (IRAS) in Bangladesh (F. Hossain et al., 2022). IRAS represented a significant technological leap in achieving higher spatial resolution from the km scale of IAS to the 30m scale using Landsat Thermal Infrared (TIR) data to estimate actual ET based on SEBAL. Combined with crop water demand from Penman-Monteith FAO 56 (Allen et al., 1998) that could already be estimated at comparable scales, IRAS offered the prediction of potential over or under-irrigation based on the comparison of actual ET estimates from SEBAL and crop water demand estimates from Penman-Monteith (Bose et al., 2021). Additional improvements in IRAS include cloud-based operationalization and automation that requires minimum internet bandwidth or local computing resources for the stakeholder agency. Climatic variables like precipitation, temperature, and wind speed are incorporated to generate advisories for farmers on how much they should water their crops based on potential over or under-irrigation in the previous week (F. Hossain et al., 2022). These advisories are then automatically transmitted to relevant users comprising farmers, irrigation district managers and water agency staff. With the combined use of two Landsat missions 8 and 9, the temporal resolution of IRAS has been improved from biweekly to weekly frequency. Additionally, dynamic crop coefficient values based on planting dates are used in IRAS to accurately estimate the actual water need for a given date.

Due to these innovative features, IRAS has been adopted by the Department of Agricultural Extension (DAE), of the Ministry of Agriculture of Bangladesh (https://iras.bamis.gov.bd/ or visit https://bamis.gov.bd and click "Satellite Products"). Currently, IRAS operates on DAE servers, providing irrigation advisories to farmers across approximately 2,000 agricultural districts located in the Northwestern and Northeastern regions of Bangladesh (Landsat Science, 2023).

Given the recent adoption of the farmer/user-centric advisory system of IRAS that optimizes water use at the consumer side, we believe there is now an opportunity for an operator-centric or water-provider-centric irrigation advisory system for surface water irrigation at the supply side. This proposed operator-centric (hereafter alternated with 'provider-centric') advisory can potentially guide irrigation water managers to optimize water delivery at the supply side where water is being centrally withdrawn from a river for dispatch to the irrigation command areas. Water suppliers and canal operators can be informed as to how much water should be allocated or delivered to each region or canal based on the water needs of all the farmers of that region. By assimilating nowcast and forecast precipitation data that are now available from weather forecast systems with the previous week's water allotment to the region and the current crop water demand, an accurate estimate of net water requirement for each region can be estimated. This net water requirement can be translated to the flux of surface water (L^3/T) that would need to be maintained in water-supplying canals at the supply side for water managers. Consequently, such a system can support dispatch decisions for surface water irrigation based on data-driven science, crop water needs, and ambient environmental conditions.

In this study, our central research question is: ***How can we dynamically optimize a multi-branched canal irrigation system using earth observation satellites?*** We believe that addressing the sustainable use of surface water requires tackling it from both the consumer side (by farmers/users) and supply side (by water managers/canal operators) with an operator-centric system as a complement to the user-centric system. The key goal of this study is to guide canal operators and help them make informed decisions on how much surface water to dispatch centrally from the source at the river for the command areas to maximize agricultural water productivity. We developed and implemented sDRIPS for the Teesta River Irrigation system located in Northern Bangladesh where there is currently an urgent need to optimize the use of surface water resources.

4.3. Study Area

The sDRIPS study focuses on the Teesta Barrage Project (TBP) in Bangladesh, the country's largest surface water irrigation project. Established in 1990, the TBP supplies irrigation water from January to April and spans the Teesta River at Dalia-Doani Point in the Lalmonirhat district (Figure 4.1 Middle Panel). This project features a 615-meter-long concrete structure equipped with 44 radial gates, providing a discharge capacity of 12,750 cubic feet per second and supporting a command area of 154,250 hectares through a 4,500-kilometer network of canals (River Research Institute, 2023). For more details on TBP, readers are referred to (River Research Institute, 2023).

The main TBP canal divides into three primary canals: Dinajpur Canal, Rangpur Canal, and Bogra Canal (Figure 4.1 Middle Panel), which further branches into an intricate network of secondary and tertiary canals. Selecting the TBP as a study site offers two primary advantages: the utilization of in-situ water supply data at the head canal and the opportunity to study a complex canal network system. A successful implementation of our study to implement sDRIPS for TBP here means it can be adapted and applied to surface water irrigation systems around the world. For this study, irrigation canal shapefiles, command areas, and in-situ water supply data were provided by the Bangladesh Water Development Board (BWDB)

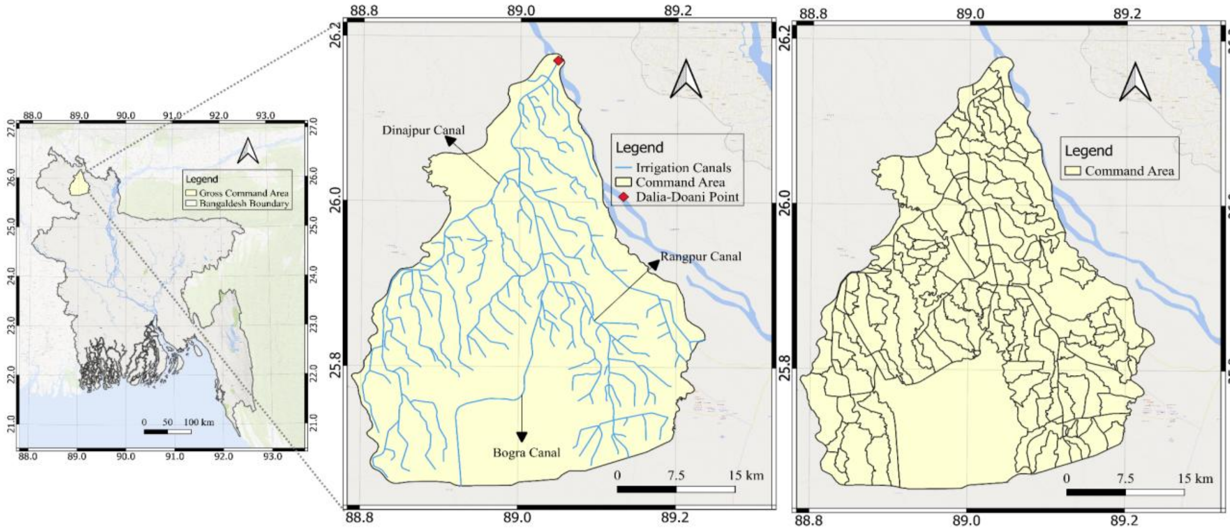

Figure 4.1: The Teesta Barrage Project (TBP). The left panel displays the map of Bangladesh with the location of TBP on the Teesta River. The middle panel illustrates the network of irrigation canals. The right panel highlights the command areas served by the primary, secondary, and tertiary canals.

4.4. Dataset Used

The publicly available datasets used in sDRIPS allow for accurate assessments of the actual water needs of command areas while considering climatic factors. To facilitate implementation and reduce technical barriers, the cloud computing platform Google Earth Engine (GEE), developed by Gorelick et al. (2017), was utilized. The cloud computing platform helps sDRIPS to eliminate the need to download and process satellite and weather data on the local machine. Table 4.1 provides a comprehensive list of the datasets utilized by sDRIPS. Some of the key datasets, which are referenced by their initials in Table 4.1 but not previously mentioned in the text are - Shuttle Radar Topography Mission (SRTM), Global Land Cover Characterization (GLCC), and National Oceanic and Atmospheric Administration (NOAA).

Table 4.1: Datasets used in sDRIPS

Dataset	Dataset ID on GEE	Band Name/Derived Products	Description	Spatial Resolution	Temporal Resolution
Global Forecasting System (GFS)	NOAA/GFS0P25	temperature_2m_above_ground	Air Temperature	25 km	6 hours
		u_component_of_wind_10m_above_ground	Wind Speed (u component)		
		u_component_of_wind_10m_above_ground	Wind Speed (v component)		
		specific_humidity_2m_above_ground	Specific Humidity		
		----	Pressure (Estimated using Hypsometric Equation)		
	GEE and NOAA	total_precipitation_surface	168-hour Precipitation		
		B2	Blue	30 m	

		B4	Red		
		B5	NIR		
Landsat 8 and Landsat 9 satellite series	LANDSAT/LC08/C02/T1_TOA and LANDSAT/LC09/C02/T1_TOA	B6	Shortwave IR		16 days individual, 8 days combined
		B10	Low Gain Thermal		
		B11	High Gain Thermal		
Sentinel 1 satellite	COPERNICUS/S1_GRD	VV	Single co-polarization, vertical transmit/vertical receive	10 m	10 days
Shuttle Radar Terrain Mapping (SRTM)	USGS/SRTMGL1_003	---	Digital Elevation Map	30 m	---
Global Land Cover Classification (GLCC)	COPERNICUS/Landcover/100m/Proba-V-C3/Global	discrete_classification	Land Use Land Cover	100 m	---
IMERG GPM	---	---	Precipitation	10 km	Around 4 hours

4.5. How sDRIPS Works

An overview of sDRIPS is first provided along with an illustration (Figure 4.2), followed by a brief description of the steps of the processes. These steps are then covered in detail in the methodology section.

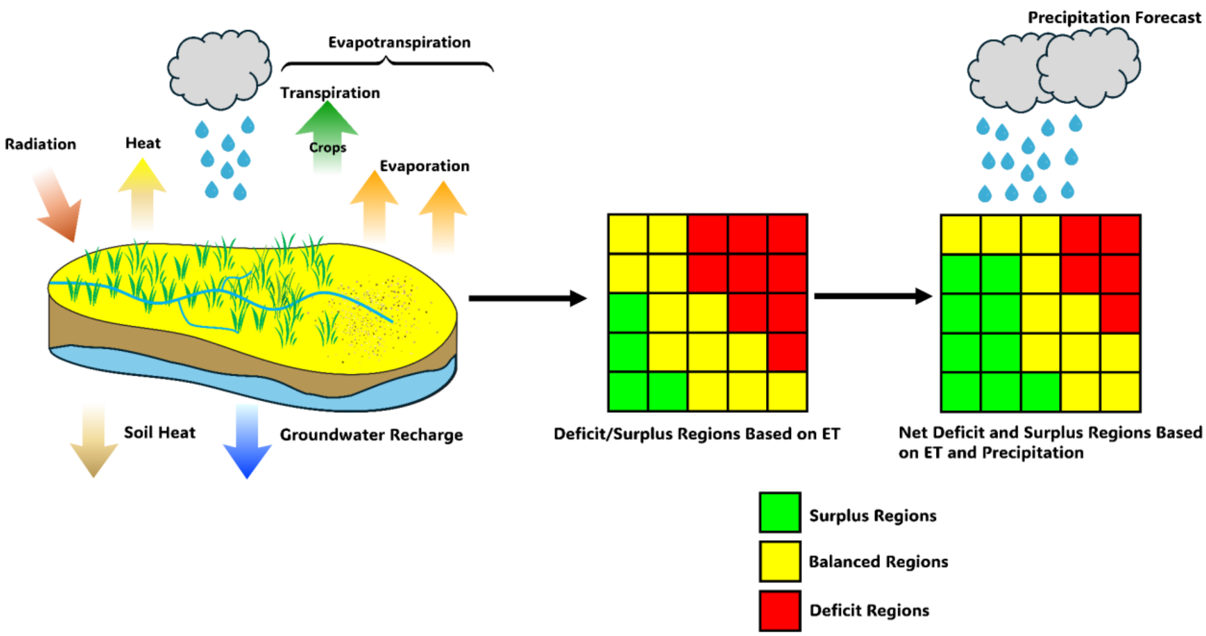

Figure 4.2: Illustration of how sDRIPS works. Green represents the regions where surplus irrigation has been provided, while red regions are deficit regions needing more irrigation.

When a Landsat satellite overpasses a command area, sDRIPS collects crucial information required to estimate the actual water needs of the crop based on its growth stage using Penman-Monteith ET, and the amount of water applied to the crop using SEBAL ET. The impact of precipitation is also considered to refine the water demand estimation. This includes both nowcast precipitation data and forecasts of incoming precipitation. The system then compares these estimates to determine whether the water needs of all crops in the given command area are met. If the net water need is not satisfied, the area is classified as a deficit region, and the volume of water needed for dispatch for the coming week is quantified (Figure 4.2). Conversely, if the water demand is met and there is an excess of provided water, the area is classified as a surplus region, The volume of surplus water that can be potentially stored locally and used later for nearby deficit regions is quantified (Figure 4.2). This quantification allows water providers to monitor and adjust the water supply and dispatch decisions from the main canal efficiently to meet the specific needs of each command area.

Figure 4.3 illustrates the proposed methodology for the irrigation advisory system, sDRIPS. Before starting, users need to have a GEE account (https://signup.earthengine.google.com/) to estimate Penman–Monteith and SEBAL evapotranspiration and a Precipitation Processing System (PPS) account (https://registration.pps.eosdis.nasa.gov/registration/) to download the GPM IMERG precipitation data.

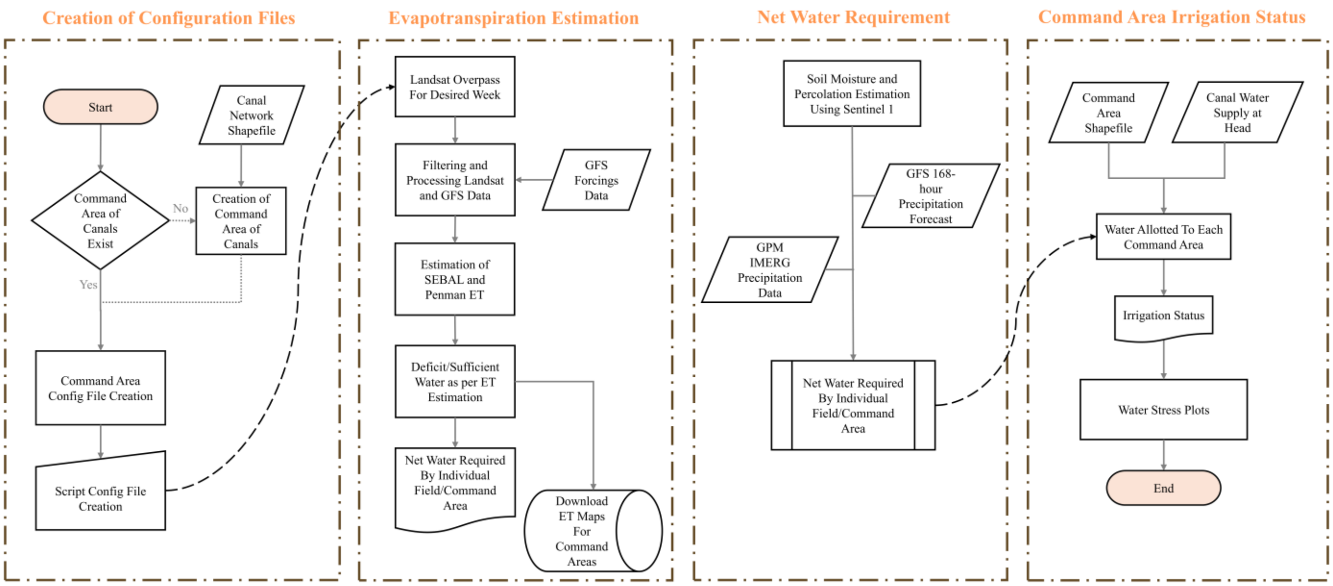

Figure 4.3: Flowchart illustrating the proposed data processing methodology for the irrigation advisory system.

Step 1: Create command areas based on the availability of command area data: This includes creating configuration files for the command areas and updating the script configuration files accordingly to run the system.

Step 2: Estimate Penman-Montieth and SEBAL ET: Use GEE to calculate evapotranspiration using the Penman-Monteith method and SEBAL.

Step 3: Estimate net water requirements for each command area: Estimating soil moisture and percolation using Sentinel 1. Download IMERG precipitation data and use it to estimate the net water requirements by combining past precipitation with GFS precipitation forecasts.

Step 4: Estimate the water supplied to each command area: Calculate the water supply at the head canal and distribute it across the command areas based on the supply.

Step 5: Assess irrigation status and generate plots: Evaluate the irrigation status using the estimated ET and water supply and generate water stress plots for analysis.

4.6. Methodology

4.6.1. Creation of Configuration Files

The configuration files provide the essential information required for the sDRIPS system to tailor the advisory system to the specific needs of the user. Readers are requested to refer to Section 8.1 of Appendix A for more details.

4.6.2. Penman-Monteith and SEBAL Based Evapotranspiration Estimation (Step 2)

After the creation of the configuration files, the next step is to estimate Penman-Monteith and SEBAL-based evapotranspiration using Landsat data. sDRIPS checks the latest Landsat imagery from both Landsat 8 and Landsat 9 collections for each command area at the specified time. Once an image for the given area is found, the timestep of the image is noted and used to estimate the Kc value by calculating the difference between the planting date and the image timestep and using the crop configuration file. Similarly, other useful information, as mentioned in Table 4.1, is stored temporarily on the cloud.

4.6.2.1. Penman-Monteith Evapotranspiration

Penman-Monteith evapotranspiration (ETo) serves as a proxy for potential water demand for a reference crop. ETo is estimated for the Landsat overpass date and for all command areas following the methodology of Allen et al., (1998). The equation for ETo is as follows:

$$ETo = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{T+273}u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \tag{4.1}$$

Where ETo is reference evapotranspiration (mm/day), R_n is net radiation at the crop surface (MJ/m²/day), G is soil heat flux density (MJ/m²/day), T is mean daily air temperature at 2 m height (°C), u_2 is wind speed at 2 m height (m/s), e_s is saturation vapor pressure (kPa), e_a is actual vapor pressure (kPa), e_s–e_a is saturation vapor pressure deficit (kPa), Δ is slope of the vapor pressure curve (kPa/°C), γ is psychrometric constant (kPa/°C).

ETo calculated using Equation (4.1 is the evapotranspiration for a reference crop. It is then adjusted for the actual crop growing in the given command area by considering the crop type, development stage, and water stress on crop transpiration for that command area using the following equation

$$ET = ETo \times K_c \times K_s \tag{4.2}$$

Where ET is the potential evapotranspiration (mm/day), K_c is the crop coefficient and K_s is the soil water stress coefficient. Unless otherwise specified, hereafter, PET will be used as a term to refer to crop water demand ET based on the Penman-Monteith equation.

4.6.2.2. SEBAL Evapotranspiration

In our study, we employed the SEBAL algorithm developed by Bastiaanssen et al. (1998). The SEBAL model has been effectively implemented in more than 30 countries (Bastiaanssen et al., 2005). The SEBAL model has demonstrated high accuracy in estimating evapotranspiration. A few examples are - Zahid et al. (2023) compared SEBAL ET with lysimeter ET in Pakistan, finding the Coefficient of Determination (R^2) of 0.9 and a Root Mean Square Error (RMSE) of 1.26 mm/day with maximum error as 24.3%. Similarly, Zoratipour et al. (2023) found an R^2 of 0.9 and an RMSE of 2.15 mm/day in Iran, and Rawat et al. (2017) validated SEBAL ET estimates in India with an R^2 of 0.91. Typically the accuracies of the estimated SEBAL ET model are 85%, 95%, and 96% at daily, seasonal, and annual scales, respectively (Tang et al., 2013). SEBAL used in the sDRIPS framework was originally employed and validated by Bose et al. (2021) and Hossain et al. (2022).

The SEBAL model solves the surface energy balance equation to estimate evapotranspiration using satellite images and meteorological forcings data. SEBAL computes an instantaneous ET flux for the Landsat overpass time. A series of equations are incorporated in the SEBAL model that computes net surface radiation, soil heat flux, and sensible heat flux to the air. The residual energy flux is then calculated by subtracting the soil and sensible heat fluxes from the net radiation at the surface. This residual energy (latent heat) enables liquid water to transition to the water vapor phase, that is, the required evapotranspiration. Thus, for each pixel of the image, the ET flux is calculated as a residual of the surface energy budget shown in Equation (4.3.

$$\lambda E = Rn - G - H \tag{4.3}$$

Where λE is latent heat flux (energy used for evapotranspiration), R_n is net radiation at the surface, G is soil heat flux, and H is sensible heat flux. SEBAL-based evapotranspiration (SEBAL ET) serves as a proxy for the actual water consumed by the crop, a concept validated by Bose et al. (2021). In our study, daily or 24-hour evapotranspiration was computed by assuming that variations in instantaneous evapotranspiration are not significant over the 24-

hour period (Allen et al., 2007). For weekly evapotranspiration calculations, we estimated evapotranspiration on the day of Landsat acquisition and considered this steady-state evapotranspiration value for the next seven days until the availability of the next Landsat image. In this study the term SEBAL ET will refer to actual ET estimated using the SEBAL method (Equation (4.3)).

4.6.2.3. Deficit Irrigation

To quantify the evapotranspiration-based water requirements from a command area, we estimate the deficit for each command area or region. The deficit was calculated using the Equation (4.4.

$$Deficit\ Irrigation\ =\ SEBAL\ ET\ -\ PenmanMonteith\ ET \qquad (4.4)$$

Figure 4.4 Panel (a) represents the cumulative sum of Penman-Monteith based ET for 7 days, depicting the total water demand based on the rice crop water requirements. Similarly, Panel (b) is the cumulative sum of SEBAL-based ET for 7 days, providing an estimate of the actual water consumed by the crops, factoring in the actual conditions of the fields. Panel (c) shows the deficit and sufficient regions. The intensity of the deficit or sufficient is estimated using Equation (4.4) for every pixel. The deficit regions indicate areas where the water demand exceeds the supplied water, while sufficient regions indicate areas where the supplied water meets the demand. This detailed mapping allows water managers to identify critical areas needing more precise water management interventions. Additionally, an uncertainty analysis of the SEBAL ET estimates was performed using Monte Carlo simulations, revealing a normalized deviation of 8.6% from the ensemble mean for the operational period of TBP. For further details, please refer to Section 8.3 of Appendix A.

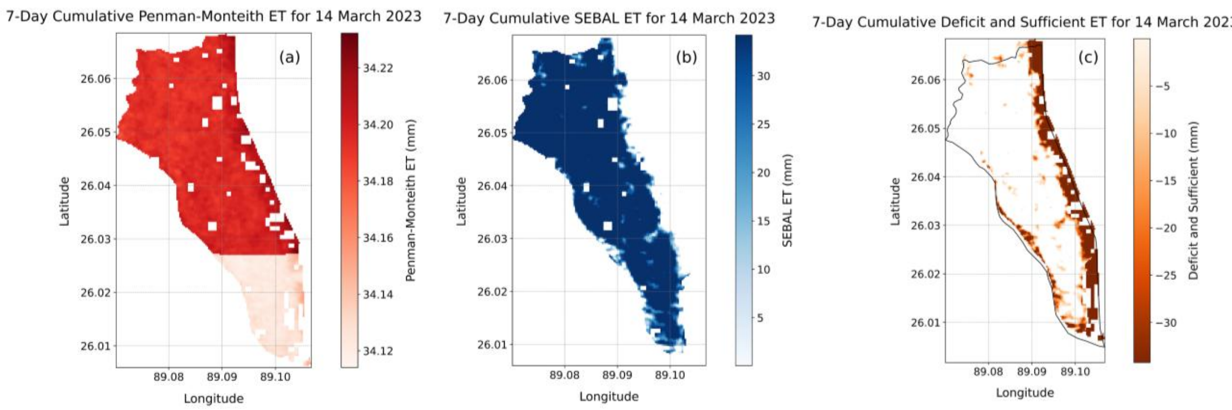

Figure 4.4: Map of estimation of ET-based deficit and sufficient irrigation regions.

4.6.3. Estimation of Net Water Requirement (Step 3)

To accurately determine the net water requirement, several factors must be considered, with soil moisture being a critical component. Soil moisture plays a pivotal role in understanding the interactions between surface and groundwater and in generating accurate irrigation advisories, as it governs the amount of water percolating through the soil. Water lost through percolation, which moves beyond the crop root zone, becomes unavailable for crops. To account for percolation in our framework, we utilized Sentinel-1 Synthetic Aperture Radar (SAR) data available on GEE. While global soil moisture products have been developed at coarser resolutions ranging from 9 km to 40 km (Chan et al., 2016; Kerr et al., 2012; Kim et al., 2023), these products are unsuitable for field-scale applications. SAR sensors, on the other hand, achieve finer spatial resolutions (10 m), making them suitable for agricultural field-scale analysis (Arias et al., 2023). Sentinel-1 C-band (wavelength of 5cm) data has been extensively used in soil moisture estimation studies and has demonstrated promising results up to 100 mm depth (Arias et al., 2023; Bauer-Marschallinger et al., 2019; Bhogapurapu et al., 2022; Wagner et al., 1999). While ground sensors provide ideal soil moisture measurements, installing them is often impractical on a large scale. To maintain the global scalability of the framework, Sentinel-1 C-band data was selected as the most viable alternative for estimating soil moisture. After estimating soil moisture from Sentinel-1 data, soil moisture at field capacity was derived using the Hengl & Gupta, (2019) dataset available on GEE. Percolation was subsequently calculated using Equation (4.5.

$$Percolation = Soil\ Moisture - Field\ Capacity \tag{4.5}$$

Figure 4.5 Panel (a) shows the average soil moisture estimated using the Sentinel 1 C band for each command area, Panel (b) shows the soil moisture at field capacity from Hengl & Gupta, (2019) and Panel (c) shows the average percolation that happened in each command area. Since percolation only happens when the soil moisture is greater than the soil moisture at field capacity, percolation was estimated using the following equation (Equation (4.6)) for each pixel and then averaged over the command area.

$$Percolation = \begin{cases} Soil\ Moisture\ - Field\ Capacity\ ; \geq 0 \\ 0;\ else \end{cases} \tag{4.6}$$

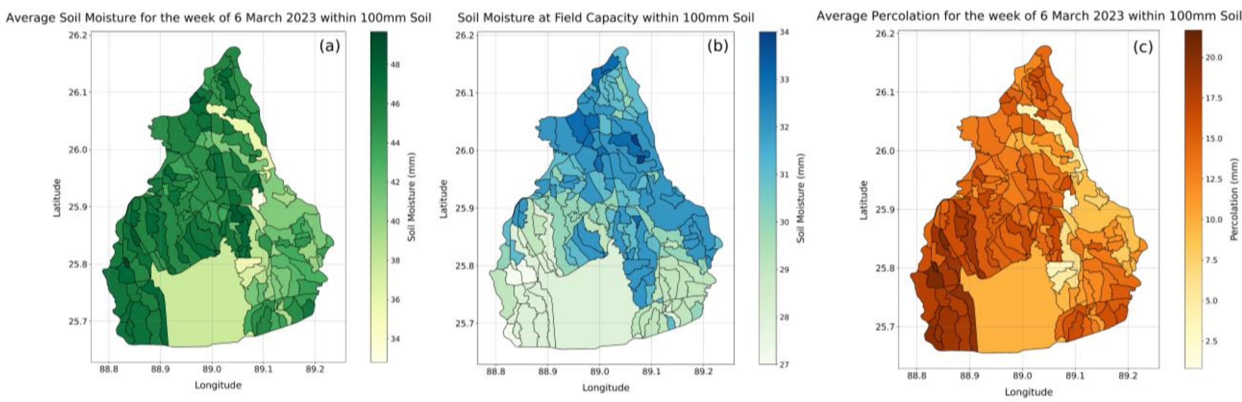

Figure 4.5: The left panel shows the average soil moisture estimated using Sentinel 1 C band on 6th March 2023, the middle panels show the soil moisture at field capacity, and the right panel shows the estimated percolation on 6th March 2023

Another crucial factor in determining the net water requirement is precipitation. Both Penman-Monteith and SEBAL evapotranspiration models do not inherently account for the effect of precipitation. To estimate the net water requirement accurately, it is essential to incorporate the precipitation events (if any). For this purpose, the precipitation for the current week or the operational week is estimated using the GPM IMERG data. Given the operational objective of the framework, only IMERG's early run data is utilized. The cumulative precipitation for the seven days preceding the latest Landsat overpass date is considered. Additionally, to provide accurate advisories to water managers, the cumulative sum of forecasted precipitation for the next seven days from the Landsat overpass date is also included in the calculations. Consequently, the net water requirement can be determined using the following equation:

$$Net\ Water\ Requirement\ = SEBAL\ ET\ -\ PET\ - Per +\ P_{nc} + P_{fc} \tag{4.7}$$

Here in Equation (4.7), SEBAL ET is the estimated evapotranspiration, PET is the evapotranspiration estimated using the Penman-Monteith (proxy of actual water demand of crop), Per is the percolation, P_{nc} is the nowcast precipitation, and P_{fc} is the forecasted precipitation.

Based on the constituents of the above equation, the net water requirement is assessed as follows:

- If the net water requirement is positive, it indicates that the combined effect of last week's water supply, the amount of percolation that happened (Equation (4.6)), the current week's precipitation, and the next week's forecasted precipitation has met or exceeded the water demand. In this case, the net water requirement for the given command area is considered to be null, as no additional water is needed in the coming week.
- If the net water requirement is negative, it signifies that the combined effect of the last week's water supply, the amount of percolation happened, the current week's precipitation, and the next week's forecasted precipitation was insufficient to meet the upcoming demand. Consequently, additional water is required for the given command area in the coming week.

It must be remembered that to calculate the surplus or deficit regions with precipitation factored in, we do not require to know the system losses, such as canal seepage, canal evaporation or sluice gate efficiency. This is because our calculations are based on the field conditions after the water has been delivered or consumed on the field. The actual ET estimated from the SEBAL algorithm is assumed to be a proxy of on-field water use or on-field irrigation.

4.6.4. Command Area Irrigation Status (Step 4-5)

After estimating the net water requirements for each command area, sDRIPS evaluates whether the current water supply conditions can meet these demands or not. To perform this evaluation, the water supply at the head canal is proportioned according to the area of each command area. This proportioning is based on the assumption that a larger command area will have more cropped land, resulting in higher evapotranspiration and, consequently, a greater water demand compared to smaller command areas. A distribution factor (DF) is created using the following equation:

$$DF_i = \frac{A_i}{\sum_{i=1}^{N} A_i} \tag{4.8}$$

Where DF_i is the distribution factor for i^{th} command area, N is the total number of command areas, A_i is an area of i^{th} command area. After estimating the distribution factor for each command area, the available water supply at the head canal is multiplied by this factor to determine the amount of water distributed to each command area. Subsequently, the amount of water passing through each canal is estimated based on the distribution to the command areas.

Figure 4.6 illustrates this distribution process, showing how water is allocated to various command areas within the irrigation canal network. In Figure 4.6, let the total water supply at the head canal be denoted as *T*. The main canal (*C*) branches into secondary canals (*S1C, S2C, and S3C*) and further into tertiary canals (*T1S2C, T1S3C, T2S3C*). The distribution factor (DF) for a command area *b* is calculated as shown in Equation (4.9):

$$DF_b = \frac{b}{a+b+c+d+e} \tag{4.9}$$

The water supplied to command area *b* is then determined by:

$$WS_b = T \times (DF_b) \tag{4.10}$$

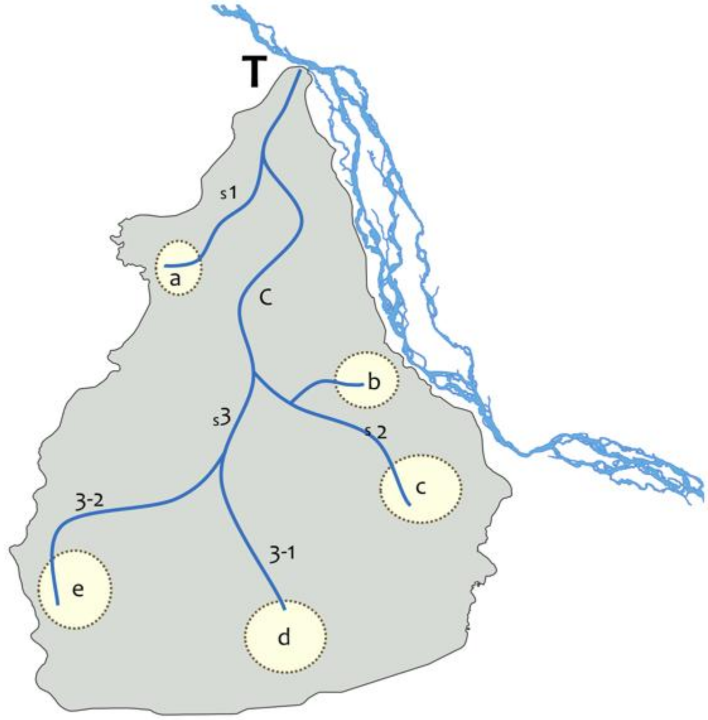

Figure 4.6: Illustration of how water distribution(dispatch) advisory for the command areas of an irrigation canal network.

Similarly, the water passing through canal *S2C* includes the amount of water supplied to its command area and the water distributed to its sub-canals (Figure 4.6). This can be expressed as below:

$$W_{S2C} = WS_b + WS_c \tag{4.11}$$

4.7. Results

4.7.1. Water Needs and Distribution in Command Areas

Figure 4.7 illustrates the four panels generated by sDRIPS after running it for 14th March 2023 for TBP. For a better understanding of the readers, these panels are discussed in detail as follows. Figure 4.7 Panel (a) depicts the main Teesta canal bifurcating into three primary canals and the percentage distribution of water in each canal. Using the DF, it is estimated that the Dinajpur canal receives approximately 27%, the Bogra canal receives around 17%, and the Rangpur canal receives about 18% of the total water supply withdrawn from the Teesta River for that week. Similarly, the water distribution for each canal is also quantified. Overall, the Teesta main Canal, carrying 100% of the water, supplies around 62% to these three primary canals, with the remaining 38% distributed to smaller canals serving command areas upstream of the primary canals. Quantifying this information is beneficial for any canal system (not limited to TBP) with some control over the distribution of water to its secondary and tertiary canals. This allows water managers to monitor and adjust the water distribution to each command area based on current needs. The percentage of water flowing in each canal was estimated using the basic principles of network theory, where the amount of water from the lowest canal with no bifurcations was summed up to the primary canal.

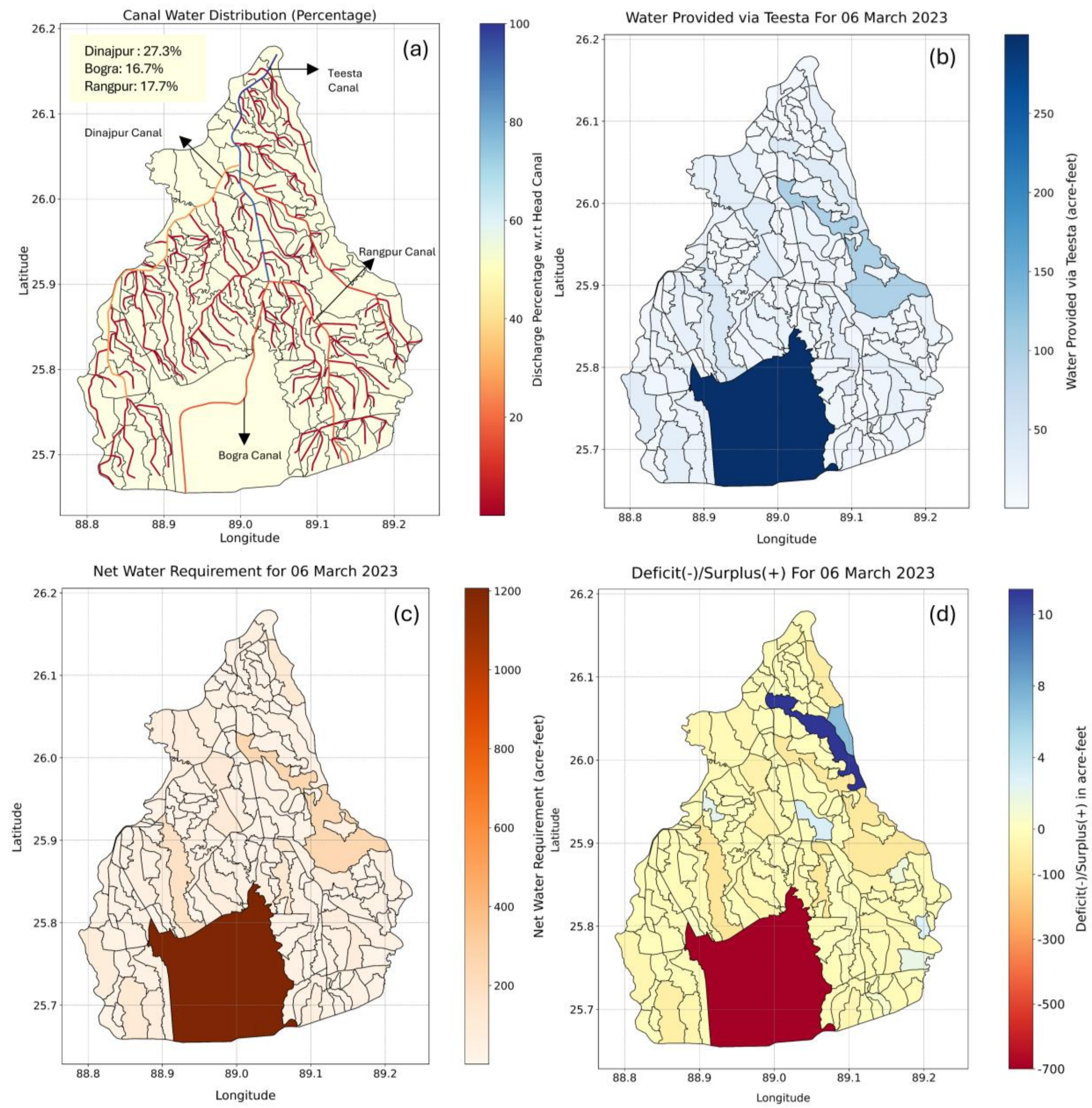

Figure 4.7: (a) Canal water distribution based on the distribution factor where head canal is assumed to have 100% water at the source; (b) amount of water allotted to each command area from the water supply based on (a); (c) net water requirement of each command area after the inclusion of previous week's water supplied to crops, nowcast, and

forecast precipitation, and amount of water percolated (d) surplus or deficit command areas after the integrating all the components of water supply (panels b and c).

Figure 4.7 Panel (b) illustrates an example of the amount of water that provided to each command area by the water supply from the Teesta River using the canal network. This is calculated by multiplying the DF with the available water supply at the head canal and following the hierarchical canal network (primary to secondary; secondary to tertiary and so on).

Figure 4.7 panel (c) illustrates the net water requirement, calculated by integrating multiple factors. These factors include the water applied to plants during the previous week, the current water needs of the plants, the precipitation received, the forecasted precipitation for the upcoming week and the amount of total water percolated. This estimation follows Equation (4.7). The combined effect of these factors, excluding the canal water supply, met the needs of 60% of the command areas for the given date. This information is crucial for the hydrologic community, including water managers and researchers, as it identifies regions dependent on precipitation variability. Over an extended period, water managers and researchers can study the areas that are primarily reliant on precipitation and are sensitive to changes in precipitation patterns. When precipitation is adequate, these regions lessen the stress on water managers. Conversely, insufficient precipitation exacerbates the stress on water resources. Additionally, these regions can vary based on crop type, growth stage of crop, and precipitation intensity. Researchers can leverage this data to study the long-term sensitivity of various crops to precipitation variability, informing crop selection in regions with high precipitation variability and recommending robust crops capable of withstanding such conditions.

Figure 4.7 panel (d) is derived by integrating the net water requirements of the command areas (Panel c) with the water provided by the available water supply (Panel b). This panel highlights regions within the TBP that have either surplus or deficit water. This information is crucial for water managers to make informed decisions about adjusting and optimizing water supply across different regions in the upcoming dispatch cycle. To provide a clear understanding to readers, we have discussed the various colored regions depicted in Panel (d) below.

Red and Blue Regions: These areas have either surplus water (blue) or deficit water (red). These regions pose a significant challenge to water managers. To address this, water managers can divert water from surplus areas to nearby deficit areas if the supply infrastructure allows it and there is buy-in from farmers. Additionally, they can also advise farmers in the deficit regions to cultivate less water-intensive crops to balance water distribution more effectively.

Light-Colored Regions: These areas exhibit low intensity of surplus or deficit water. Water managers should aim to stabilize the red and blue regions to fall within this range. These regions typically include areas where - i) The water needs were already met by precipitation, and additional water was supplied. ii) The water needs were only partially met by precipitation, but the combined effect of the water supply sufficiently met the requirements, albeit with a slight surplus.

4.7.2. Reliability of sDRIPS

To promote the applicability of sDRIPS in other regions where the proposed framework might be suitable and to instill confidence in its reliability among readers, we conducted a two-pronged investigation: i) comparing the estimates of sDRIPS with in-situ data, and ii) comparing the estimates of sDRIPS with OpenET. In the first approach, we evaluated the net water requirement for the entire TBP by aggregating the net water requirement estimates for each command area. These estimates were calculated based on the combined response of the previous irrigation cycle, nowcast and forecast precipitation, and the amount of water percolated, as outlined in Equation (4.7). For consistency in the analysis, we assumed rice, a water-intensive crop, was cultivated uniformly across all regions of the TBP. We then compared the sDRIPS-derived net water requirement with the actual amount of water supplied through the main canal system of the TBP during the period from January 2023 to April 2023. Additionally, an uncertainty analysis was performed on the net water requirement by introducing perturbations in the parameters of the evapotranspiration models and running Monte Carlo simulations. For more details, readers are referred to Section 8.3 of the Appendix A.

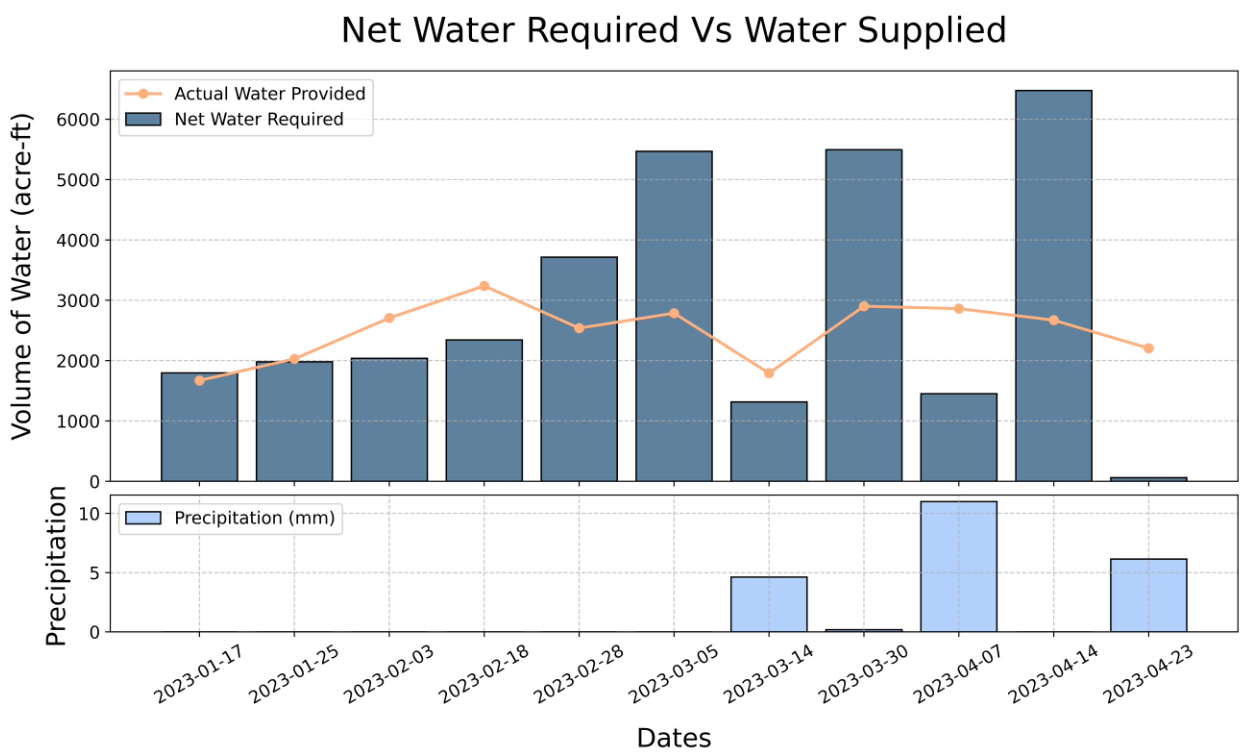

Figure 4.8: Timeseries of net water required by TBP vs amount of water supplied to TBP.

From Figure 4.8, we can infer the following:

1. Early Crop Stage: At the early stage of crop growth, the water supplied through the TBP canal sufficiently met the crop's water requirements. However, during this period, there can be instances where the water supplied exceeded the crop water needs, indicating potential inefficiencies in water allocation.

2. Mid-Crop Stage: During the mid-stage, the water supply from the TBP canal was insufficient to meet the crop water requirements, resulting in unmet water demand.

3. Precipitation Effects: Precipitation intensity significantly reduced the water burden on the TBP canal during certain weeks:

    a. 14th March 2023, 7th April 2023, and 23rd April 2023: Precipitation alone satisfied the crop water needs for the TBP region, demonstrating that the TBP region could have significantly benefitted from integrating precipitation and its forecasting into its water allocation planning. This could reduce unnecessary reliance on canal water during times of sufficient rainfall.

    b. 30th March 2023: Despite the precipitation, the combined effect of rainfall and TBP water supply was inadequate to meet crop water demands due to low precipitation intensity.

4. Late Crop Stage: Toward the end of the crop cycle, both water demand and TBP water supply decreased. During this period, precipitation also occurred, further reducing the canal's burden.

In the second approach, we compared the sDRIPS estimates (SEBAL ET as implemented within the sDRIPS framework) with the OpenET. OpenET leverages 30m Landsat imagery and employs six state-of-the-art satellite based energy balance models, that is, ALEXI/DisALEXI, eeMETRIC, geeSEBAL, PT-JPL, SIMS and SSEBop (details on these models can be found in Volk et al., (2024)). These models have been extensively applied and evaluated in the United States for various water management and agricultural applications (Volk et al., 2024). According to Volk et al., (2024), a comparison of OpenET data with 152 in-situ datasets across the USA showed that the daily ensemble results for cropland sites had a mean absolute error of 23.6% and an RMSE of 31.1% of the mean. However, OpenET is currently limited to use within the USA. To enable a fair comparison, we applied the sDRIPS framework to regions within the USA where OpenET data is available. Specifically, we selected two locations for comparison, as illustrated in Figure 4.9 and Figure 4.10

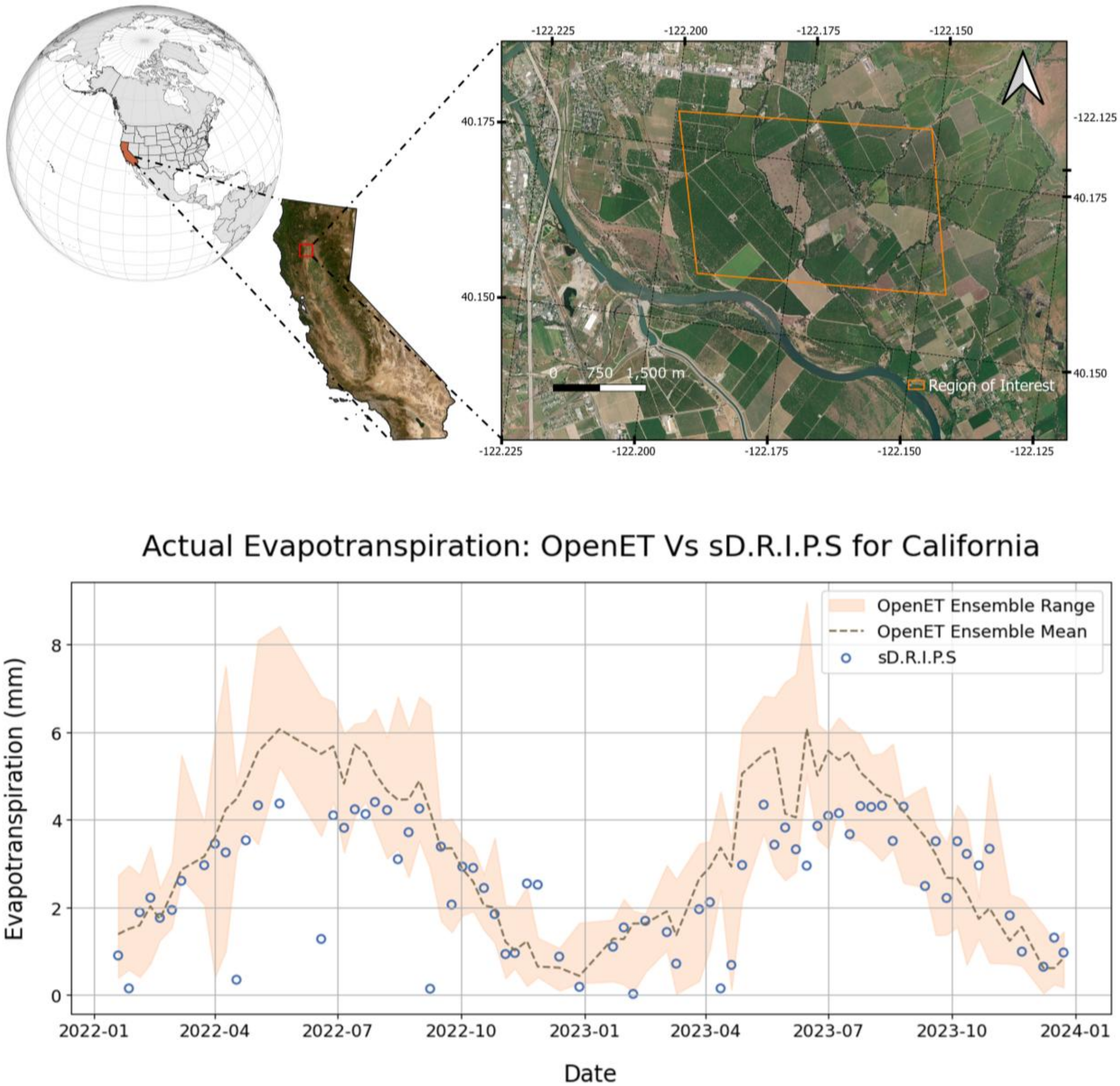

Figure 4.9: Location of the region of interest in California with the timeseries of estimates from OpenET and sDRIPS-based SEBAL ET.

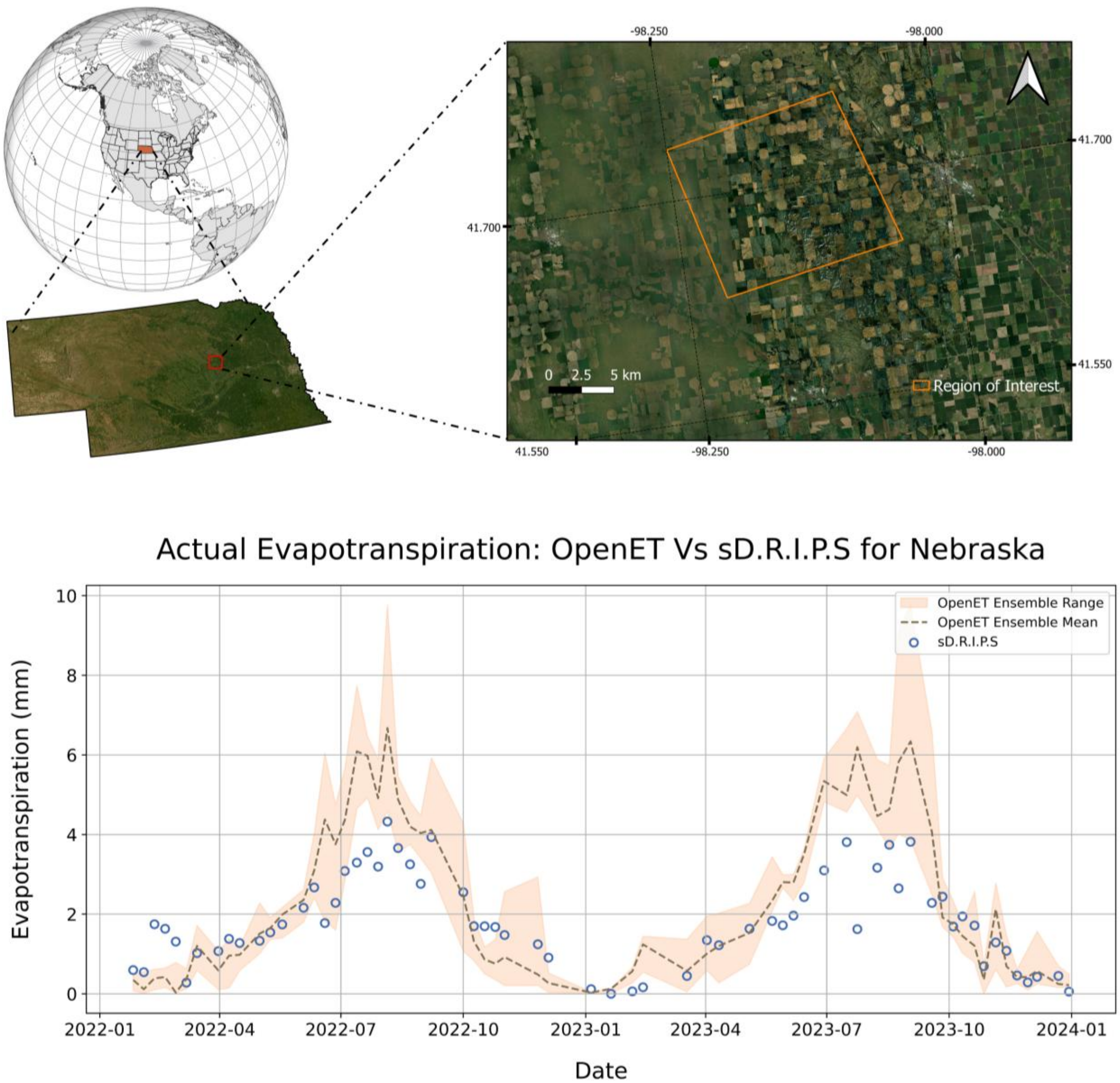

Figure 4.10: Location of the region of interest in Nebraska with the timeseries of estimates from OpenET and sDRIPS

In both Figure 4.9 and Figure 4.10, the upper panel shows the location of the region of interest where sDRIPS was applied, and the lower panel shows the timeseries of the estimates. The sDRIPS estimates align closely with the OpenET ensemble range, although some outliers are observed. Excluding the outliers, two key inferences can be drawn from the comparison:

1. Both sDRIPS and OpenET exhibit similar trends in ET estimates across the studied regions.
2. During peak periods (e.g., July), sDRIPS estimates slightly underpredict ET, or OpenET slightly overpredicts it (noting that OpenET itself is not 100% accurate). This discrepancy is likely due to differences in the spatial resolution of meteorological input data that comes from different sources. OpenET utilizes Gridded Surface Meteorological (gridMET)

(Abatzoglou, 2013) and North American Land Data Assimilation System (NLDAS) products (Xia et al., 2012). Both gridMET and NLDAS are limited to the USA with spatial resolution of 4 km and around 13 km respectively; whereas sDRIPS relies on a global (GFS) product with a coarser spatial resolution of 25 km to obtain the meteorological variables (see Table 4.1).

4.8. Comparative Assessment of sDRIPS Scenarios

In this section, we perform a comparative assessment of sDRIPS with the help of two scenarios. In scenario one, we simulate traditional irrigation practices, whereas in scenario two, we simulate scenario one using sDRIPS. More details are provided in the following subsections.

4.8.1. First Scenario - Traditional Irrigation Practices (without sDRIPS)

In this scenario, we illustrate traditional irrigation practices from the perspectives of both water providers and water consumers. The traditional practice from the water provider's perspective involves allotting water without knowledge of the specific needs of command areas and without accounting for precipitation. From the water consumers' perspective, traditional irrigation involves watering crops based solely on the water allocated to them, again without considering precipitation. It is important to remember that the TBP operates mostly from January to April, prior to the next monsoon season. During this period, farmers do not account for precipitation events in their irrigation practices. We illustrate the scenario simulation using the panels shown in Figure 4.11.

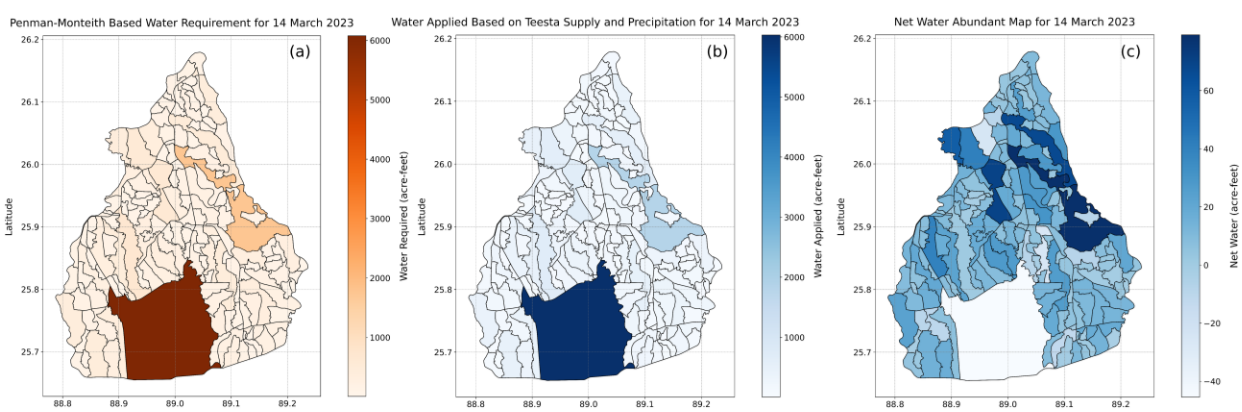

Figure 4.11: Scenario 1 example without the s.D.R.I.P.S - Simulation of traditional irrigation practices i.e. without understanding and quantifying the actual water need of the crop and command area.

Panel (a) of Figure 4.11 shows the water requirements estimated for each command area for rice crops using the FAO-recommended Penman-Monteith equation. Panel (b) of

Figure 4.11 depicts the amount of water applied to each command area by farmers (for the week of March 14, 2023), including water added from precipitation events. However, farmers do not anticipate these precipitation events and fully utilize the allotted water, disregarding previous irrigation cycle and precipitation. Panel (c) of Figure 4.11 illustrates the excess water present in the fields, calculated as the difference between the total water applied to the command area and the Penman-Monteith based water requirements of the crops. The positive values in Panel (c) of Figure 4.11 indicate that excess water has been used, which is unsustainable as it can wash away vital nutrients needed by crops, ultimately affecting yield.

4.8.2. Second Scenario - Adaptive Irrigation Practice (With sDRIPS)

In this scenario, we consider the same time frame as in Scenario 1 and for the same week; however, here, water providers use the sDRIPS system. This system accounts for water applied in the previous irrigation cycle (and whether it was sufficient or deficit), current precipitation events, and forecasted precipitation. Water is allocated based on the actual needs of the command areas, considering these factors. The simulation is illustrated with the panels shown in Figure 4.12.

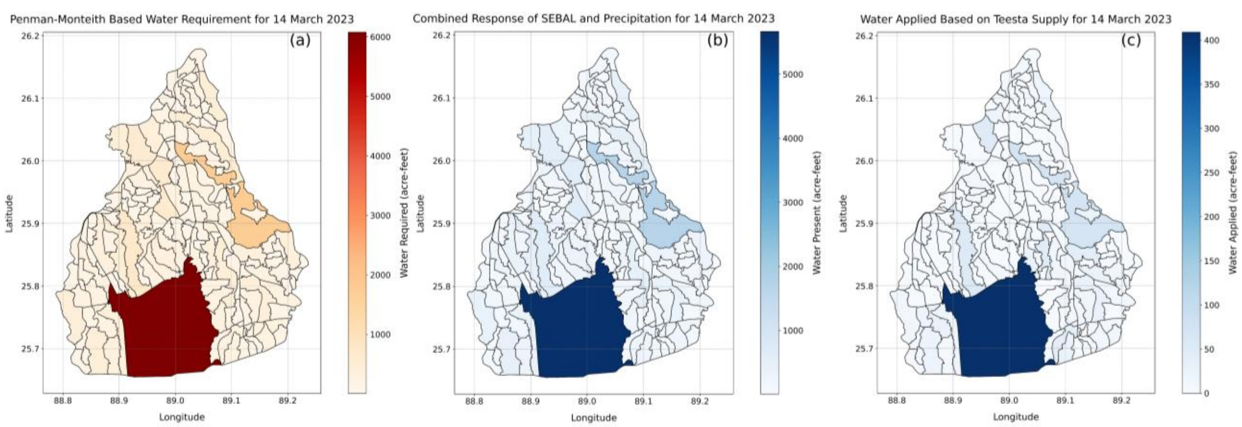

Figure 4.12: Illustration of Scenario 2 with the s.D.R.I.P.S - Simulation of adapting irrigation practices i.e. understanding and quantifying the actual water need of the crop and command area.

Panel (a) of Figure 4.12 is similar to Panel (a) of Figure 4.12 and is estimated using the Penman-Monteith equation. Panel (b) of Figure 4.12 shows the combined response, considering the previous irrigation cycle, the recent precipitation event, and forecasted precipitation before the next irrigation cycle. Including these factors reduces the stress on water providers and guides them to allocate water only where it is needed. Panel (c) of Figure 4.12 illustrates the amount of water applied by the farmers. The water was allocated to each command area specific to its needs and was fully utilized by the farmers.

By comparing Scenario 1 (traditional irrigation practices) and Scenario 2 (optimized irrigation with sDRIPS), approximately 475 acre-feet of water could have been saved using sDRIPS for the specific week as an example. The saved water could then be diverted to water deficit regions outside the TBP or stored for future TBP needs. This comparison focused on March 14, 2023, when the available water supply exceeded the water demand of the region. However, there are instances where the net demand exceeds the total water supply, as illustrated in the panels of Figure 4.13 for February 25, 2023.

Panel (a) of Figure 4.13 shows the Penman-Monteith based water needs for each command area. Panel (b) of Figure 4.13 depicts the combined response, including the effects of the previous irrigation cycle and precipitation. Panel (c) of Figure 4.13 illustrates the water applied to the command areas with the limited water supply. Panel (d) of Figure 4.13 presents the water stress map, highlighting regions where water demands were not met. In situations where the water demand exceeds the supply, if water providers or canal operators have greater control over the irrigation canals, they should reconsider the DF of the command areas. This would allow for the redistribution of water from surplus regions to deficit regions. If such control is not feasible, water providers should either manage the water needs by diverting water from outside the TBP to required command areas or encourage farmers to grow less water intensive crops in those regions.

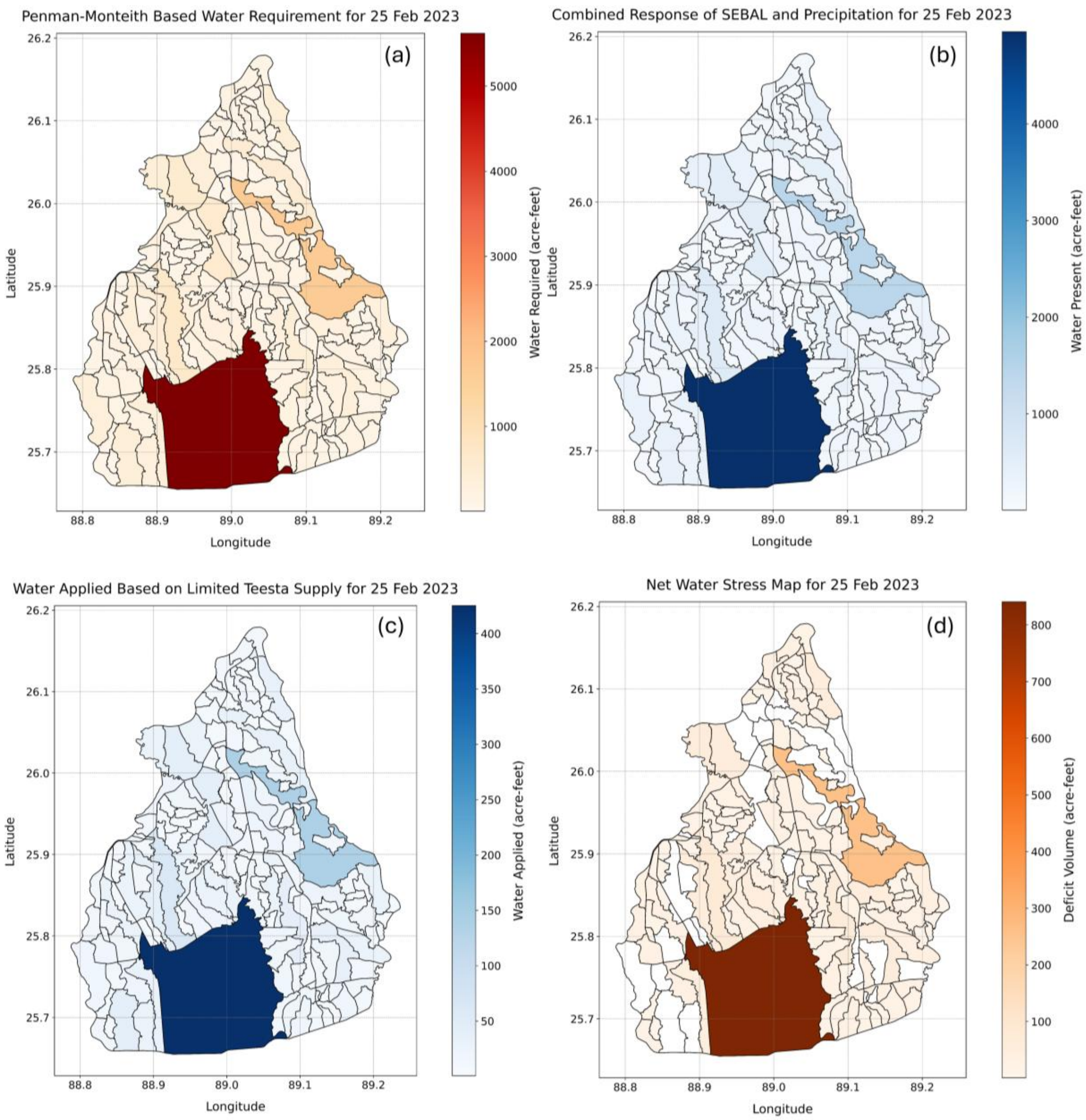

Figure 4.13: sDRIPS simulation for 25th March 2023, when the net water demand was higher than the water supply.

4.9. User Friendliness of sDRIPS and Limitations

4.9.1. User Friendliness

The sDRIPS system is designed to address the gaps identified in the current literature, as mentioned in section 4.2 of the study, by providing a user-friendly and customizable tool for analyzing various irrigation scenarios. These are discussed in detail in Section 8.2 of Appendix A.

4.9.2. Limitations

While sDRIPS provides a scalable and data-driven framework for surface water irrigation optimization, it is not without limitations. These limitations arise from sensor constraints, practical assumptions required for scalability, and certain challenges that lie outside the scope of this study. The following key limitations are acknowledged:

1. sDRIPS relies on optical sensors - Landsat series, to estimate actual ET using SEBAL. High cloud cover, common in certain regions or seasons, can compromise the framework's effectiveness by hindering the energy balance.

2. sDRIPS identifies crop fields using the Copernicus Global Land Cover dataset, though land cover maps may not always fully capture dynamic changes over time, potentially affecting net water requirement estimates. However, the framework is designed to be flexible, allowing users to refine or integrate higher-resolution local land cover data, when available, to better distinguish between irrigated and rainfed areas for improved water management.

3. Groundwater contributed minimally or negligibly to crop evapotranspiration during the study period due to the considerable depth of the water table relative to crop root lengths. However, in regions with shallower water tables or highly conductive soils, groundwater may influence evapotranspiration estimates. Current state-of-the-art satellite-based techniques often lack the spatial resolution needed to precisely quantify groundwater contributions at the farm scale. Integrating model-based simulations or in situ groundwater flow measurements into the framework's water balance equations could enhance estimation accuracy and provide a more comprehensive understanding of water dynamics.

4. The framework assumes stable climatic conditions over a weekly timescale to estimate the cumulative ET. While this assumption aligns with regional climatology in the study area and is supported by other studies (Bose et al., 2021), it may not hold true in regions where climatology exhibits high variability within a week. Consequently, users should carefully evaluate the applicability of this assumption when deploying sDRIPS in areas with high dynamic climatic conditions.

5. The framework does not account for the behavior of water-providers and water-consumers, which can significantly influence the effectiveness of any irrigation planning scheme. Behavioral change, particularly among farmers, is a gradual and long-term process that falls outside the direct scope of this study. While this study demonstrates the utility of a satellite data-driven surface water irrigation optimization tool, stakeholder-led interventions, such as economic incentives (e.g., price differentiation), may be necessary to accelerate behavior change, as discussed by Portoghese et al., (2021). Furthermore, behavioral dynamics can vary across different regions and countries, necessitating tailored approaches for each context.

6. Hydraulic constraints are assumed to be invariant in the study. Users can integrate the sDRIPS framework with different hydraulic designs and can check the flexibility (see Sawassi et al., 2022).
7. The assumption that water percolation below the 100 mm depth is lost to deep percolation may lead to an overestimation of irrigation needs, especially for crops that are known to utilize water beyond this depth. This simplification is a conservative approach that prioritizes caution in estimating crop water requirements, which is particularly important in water-scarce regions where slight overestimating water needs could lead to more effective water management decisions. However, it also introduces uncertainty, and future work could refine this aspect by incorporating models that predict root zone soil moisture or by using satellite-based soil moisture estimates for deeper layers. Additionally, methodologies such as those proposed by (Baldwin et al., 2017) could be explored to improve the estimation of water availability in deeper soil layers.

4.10. Conclusion

In this study, we co-developed and implemented a water-provider-centric irrigation advisory system, sDRIPS, to manage water resources sustainably and robustly while addressing existing gaps in the literature. sDRIPS integrates satellite and model data to estimate the actual water needs of command areas. This crucial information aids water providers in determining the appropriate water allocation for each command area. The actual water need depends on several factors, including the amount of water applied in the previous irrigation cycle, current precipitation, and forecasted precipitation. By considering these factors and the available water supply, sDRIPS advises canal operators on which command areas require water and the amount needed. During periods of excess water supply, sDRIPS advises canal operators on how much water can be stored for future irrigation needs or diverted to water-deficit regions outside the irrigation project. Conversely, during water shortages, sDRIPS generates water stress maps to identify regions experiencing water shortages and guides on the amount of water needed in deficit areas. sDRIPS also allows for the simulation of various scenarios using historical data on different crops. This capability enables stakeholders to evaluate the impact of future policies on water supply conditions if these policies are implemented today and to devise science-based responses. By offering a flexible and comprehensive tool for water management, sDRIPS has the potential to contribute significantly to sustainable irrigation practices globally, addressing the challenges posed by a changing climate and inefficiency in managing surface water and groundwater resources.

The work presented in Chapter 4 centered on a water-provider centric perspective of irrigation management, focusing on canal operators and their role in optimizing water distribution. A complementary perspective is the water-consumer centric approach, which emphasizes farmers' decision-making and water use at the field scale. While this complementary approach has been examined separately in stakeholder-specific case studies, integrating them within a unified framework can provide a more holistic understanding of water management. However, both water providers and water consumers operate under highly diverse agricultural, climatic, and technological conditions. There is no single, monolithic agricultural setting as each region varies in its infrastructure, data availability, and technical capacity. Therefore, developing a unified framework that can accommodate this diversity while addressing the needs of both types of users is essential for promoting sustainable and equitable water allocation practices. To ensure global accessibility, such a framework must also lower the barrier to entry by leveraging publicly available datasets, open-source tools, and comprehensive documentation. Chapter 5 introduces a standalone, open-source Python package that embodies these principles, integrating both provider and consumer perspectives within a single, adaptable environment.

Chapter 5. sDRIPS: A Cloud-Based, Open-Source Python Package for Satellite-Informed Surface Water Irrigation Optimization

5.1. Abstract

Surface-water irrigation underpins global food production, yet its management is increasingly strained by climate variability. Shifts in temperature and precipitation patterns create uncertainty for farmers and water managers, often causing overwatering or underwatering, which reduces crop yields and limits available water resources. Current irrigation advisory systems exhibit key limitations: most are not open-source, operate at coarse spatio-temporal scales, lack scalability, and offer limited integration of local inputs such as in situ sensors or station-based weather data. Advances in satellite remote sensing, open-access datasets, and cloud computing offer opportunities to overcome these gaps. Here we introduce sDRIPS (satellite Data Rendered Irrigation using Penman-Monteith and SEBAL), a modular, open-source Python package that generates weekly, locally tailored irrigation advisories at the farm scale (30m). sDRIPS integrates satellite observations, global weather model products, and ground-based sensor measurements when available. Its flexible architecture accommodates diverse agricultural contexts, enabling estimation of evapotranspiration, incorporation of nowcast and forecasted precipitation, and computation of net irrigation demand at field and regional scales. To lower adoption barriers, sDRIPS includes tutorials and documentation. By translating state-of-the-art Earth observation and model data into an accessible, customizable decision-support system, sDRIPS enables optimized surface-water irrigation planning under growing climate uncertainty and water scarcity.

Note: This chapter has been submitted in its current form as an article *Digital Water*. The only differences are in the numbering of figures, tables, and references to other sections.

5.2. Introduction

5.2.1. Background

Surface water irrigation remains a cornerstone of global food security, particularly in agrarian economies where canal systems, lakes, and reservoirs serve as the primary means of water delivery. However, the effective management of these systems is increasingly threatened by climate variability. Rising temperatures, erratic precipitation patterns, and shifts in seasonal water availability introduce significant uncertainty for both farmers and irrigation managers. Many of the stakeholders and farmers lack the tools to incorporate short-term precipitation forecasts or estimate water demand at the command area level of the irrigation system. This leads to inefficient water distribution practices such as overwatering or underwatering, both of which are detrimental to crop yield and water-use efficiency (F. Hossain et al., 2017). In addition to the direct impacts of climate change on crop production, there is a growing concern about future agricultural water requirements vis-a-vis water availability under the combined effects of climate change, growing population demands, and competition from other economic sectors under future socio-economic development (Calzadilla et al., 2013; Fischer et al., 2007). Addressing these multi-scalar challenges requires decision support systems that can simultaneously optimize water allocations and ensure sustainability.

Current irrigation advisory systems face several critical limitations. Many rely on dense networks of sensors and data loggers (Forcén-Muñoz et al., 2022), which are often cost-prohibitive and logistically challenging to deploy across large or resource-constrained regions. Commercial platforms are frequently subscription-based and tailored to specific regions, limiting scalability, especially in developing countries. Moreover, many open-source tools operate at coarse spatial and temporal resolutions (Chen et al., 2021; Smith, 1992; Tasumi, 2019) and lack integration with local datasets, such as in-situ weather stations or farm-level sensors. Although these technical shortcomings are well documented in literature, particularly from the perspective of water users (i.e., farmers), they remain largely unresolved in practice (Karthikeyan et al., 2020; Sishodia et al., 2020). Compounding to this issue, there is a relative absence of research and tools on the challenges faced by water providers (canal operators). These stakeholders often struggle with issues such as unreliable water supply due to upstream usage, transboundary water disputes, or infrastructure limitations.

Recent advances in satellite remote sensing and cloud-based computing offer a promising pathway to address these limitations. Freely available Earth observation datasets provide global coverage, improved temporal frequency (by combining multiple satellites), and reasonable accuracy, at no cost. Platforms such as Google Earth Engine (GEE) allow researchers and practitioners to process large volumes of satellite data without the need for extensive local computing infrastructure (Gorelick et al., 2017). Despite this progress, there remains a notable gap: few open-source, operational tools exist that harness satellite data into actionable irrigation advisories for both water users and water managers, of which, to the best of our knowledge, none of them are scalable and customizable based on the user's needs.

To address this gap, we present ***sDRIPS (s**atellite **D**ata **R**endered **I**rrigation using **P**enman-Monteith and **S**EBAL,* pronounced as 'drips'***)*** - a Python-based, open-source tool designed to generate timely, data-driven irrigation advisories at both farm and command-area scales (based on user needs). By leveraging Earth observation data and evapotranspiration models, sDRIPS aims to democratize access to satellite-based irrigation and foster more equitable, efficient water use under changing climatic conditions.

5.2.2. sDRIPS

sDRIPS is a cloud-based, open-source irrigation advisory system designed to optimize surface water use by integrating Earth observation datasets and numerical weather models. The system harnesses publicly available data from the Landsat satellite series and Sentinel 1, the Global Forecast System (GFS), and the Global Precipitation Measurement's (GPM) IMERG data (precipitation dataset, Huffman et al., 2015). At its core, sDRIPS estimates evapotranspiration (ET) using two models - the Penman-Monteith method (Allen et al., 1998) as a proxy for crop water demand under ideal (non-stressed) conditions, and the Surface Energy Balance Algorithm for Land (SEBAL) (Bastiaanssen et al., 1998) to estimate actual ET under prevailing field conditions. An underlying concept used in multiple studies (Bose et al., 2021; F. Hossain et al., 2017, 2022) that led to the evolution of the current version of sDRIPS. Building on these ET estimates, sDRIPS integrates nowcast and forecasted precipitation data, as well as estimates of percolation, to compute a net irrigation requirement. By applying both energy and water balance principles, the system generates actionable irrigation advisories at multiple spatial scales - from individual farms to entire command areas. For more detailed information, readers are advised to refer to https://sdrips.readthedocs.io/en/latest/.

5.3. The Need For a Cloud-based Open-Source Python Package

Over the years, sDRIPS has undergone several iterations, each tailored to address the diverse challenges faced by distinct user groups in diverse geographies and contexts. These customized versions were independently developed to meet the unique needs of farmers, landscape workers, and irrigation authorities. For example:

1. *sDRIPS – TBP* was developed for the Teesta Barrage Irrigation Project (TBP) in Bangladesh, providing irrigation advisories to canal operators by calculating net water

requirements at the command area level and allocating river inflow proportionally among the canals based on net water demand and the supply.

2. *sDRIPS – Sense*, deployed at the University of Washington (Seattle, USA), integrates satellite observations with in-situ sensor data to guide campus landscape irrigation. The system automatically switches between a "satellite-only" mode and a "satellite + sensor" mode, depending on the availability of local sensor data, helping to improve accuracy in meteorological inputs, especially air temperature, which can be biased in coarse-resolution global models (Smarter Irrigation for a Greener UW | UW-CEE, 2024).
3. *sDRIPS – Ukulima* used in Crown Farms in South Africa, generates irrigation advisories to farmers depending on the crop grown in each field and has the ability to integrate local weather station data to support accurate, field-scale water management (Zawya News, 2024)

While these separate deployments have demonstrated the adaptability and utility of sDRIPS, their development also revealed a critical need for *a single, centralized version of sDRIPS that integrates all prior capabilities into one modular, and flexible system*. Rather than maintaining multiple customized versions for different regions and users, a unified Python package now allows any user, whether a water provider (canal operators) or water consumer (farmer), to configure and run the tool to meet their specific requirements.

This centralized architecture is expected to significantly reduce duplication of effort, simplify maintenance, enhancements, and lower the barrier to entry for first-time users. Users no longer need bespoke versions of the sDRIPS; instead, they can tailor the centralized sDRIPS package through configuration files, optional modules, and different data integration modes. The development of this open-source, cloud-compatible Python package was further motivated by increasing global interest in the tool, following high-visibility coverage (*Satellite Data Helps Bangladeshi Farmers Save Water, Money, Energy | Landsat Science*, 2023). To respond to this growing demand, the goal of this study was to democratize the use of publicly available Earth observation data and numerical weather models by packaging sDRIPS into a freely accessible, extensible software tool for the agricultural water management community.

Recognizing the heterogeneity of the end users who differ in their technical expertise, operational scales, field conditions, and data availability, the sDRIPS package has been designed to be both user-friendly and adaptable. This package is equipped with comprehensive documentation hosted on ReadTheDocs (https://sdrips.readthedocs.io/), including Jupyter Notebook tutorials and Command-Line Interface (CLI) guides. As an open-source project hosted on GitHub (https://github.com/UW-SASWE/sDRIPS), it encourages contributions from the global community of researchers and practitioners.

By providing a centralized Python package, the sDRIPS enables diverse stakeholders to generate irrigation advisories that are tailored, scalable, and science-driven for their region of interest. In doing so, the centralized package contributes to more equitable and sustainable water management practices, that is particularly critical in an era of increasing climate variability, resource scarcity, and food security challenges.

5.4. Key Innovations of the sDRIPS Python Package

Prior to the development of the sDRIPS Python package, the system existed as fragmented and region-specific scripts for irrigation advisories – for example, Irrigation Advisory System (IAS, Hossain et al., (2017)), Provision for Advisory on Necessary Irrigation (PANI, Hossain et al., (2020)), Integrated Rice Advisory System (IRAS, Hossain et al., (2022)), sDRIPS-TBP, sDRIPS-Sense, sDRIPS-Ukulima. These implementations were manually duplicated and adapted for each new setting, making the system difficult to maintain, prone to errors, and challenging to scale. The formalization of sDRIPS into a modular, centralized Python package represents a significant leap forward in terms of usability, scalability, reproducibility, and impact. The key innovations introduced through this packaging effort include:

1. **Unified, open-source architecture integrating all functionalities** – The sDRIPS Python package consolidates all prior region-specific versions into a single, centralized platform that leverages GEE, freely available Earth observation datasets, and global weather model data. This integration provides a consistent and streamlined framework for generating weekly irrigation advisories, regardless of geographic location or user type (e.g., farmers or water managers). By bringing all functionalities "under one roof", sDRIPS ensures consistency in implementation while preserving flexibility through configuration-file-based customization.

2. **Cross platform compatibility and streamlined installation** – Recognizing the technical barriers faced by many users, sDRIPS has been designed for easy installation across major operating systems (OS) (Windows, macOS, and Linux). The package minimizes the burden of managing dependencies and environment setup challenges that often hinder the adoption of scientific software tools, thereby facilitating broader uptake by researchers, practitioners, and decision makers.

3. **Scalability through parallel processing** – To support operational use across large-scale irrigation systems or farm networks, sDRIPS incorporates parallel processing capabilities. This allows users to run analyses across multiple farm plots or command areas simultaneously, significantly reducing computation time.

4. **Robust Error Handling and Logging Mechanisms** – sDRIPS implements a "fail-fast and fail-safe" philosophy, ensuring that potential errors are caught early and clearly reported through logging. This idea is often underappreciated in hydrological tool development. It prevents time-consuming failures during long computational runs and improves debugging efficiency.

5. **Scenario simulation capability** – One of the unique features of sDRIPS is its ability to simulate hypothetical "what-if" scenarios related to crop selection, water allocation, and policy interventions. By allowing users to evaluate the impacts of future decisions using current and historical weather and water availability data, the package becomes not only a tool for operationalized irrigation management but also a platform for long-term strategic planning and research purposes. This functionality is particularly valuable for irrigators, canal operators, and policymakers aiming to assess supply-demand dynamics under various management strategies or climate conditions.

6. **Transparency and reproducibility** – Hosted on GitHub, sDRIPS adheres to the principles of open science by enabling full version control, collaborative development, and transparent issue tracking. This supports reproducible workflows and can lead to community-driven innovation.

Figure 5.1 illustrates the key software design advancements introduced in the sDRIPS package, while Figure 5.2 depicts the technological evolution of the irrigation advisory system that contributed to the development of sDRIPS.

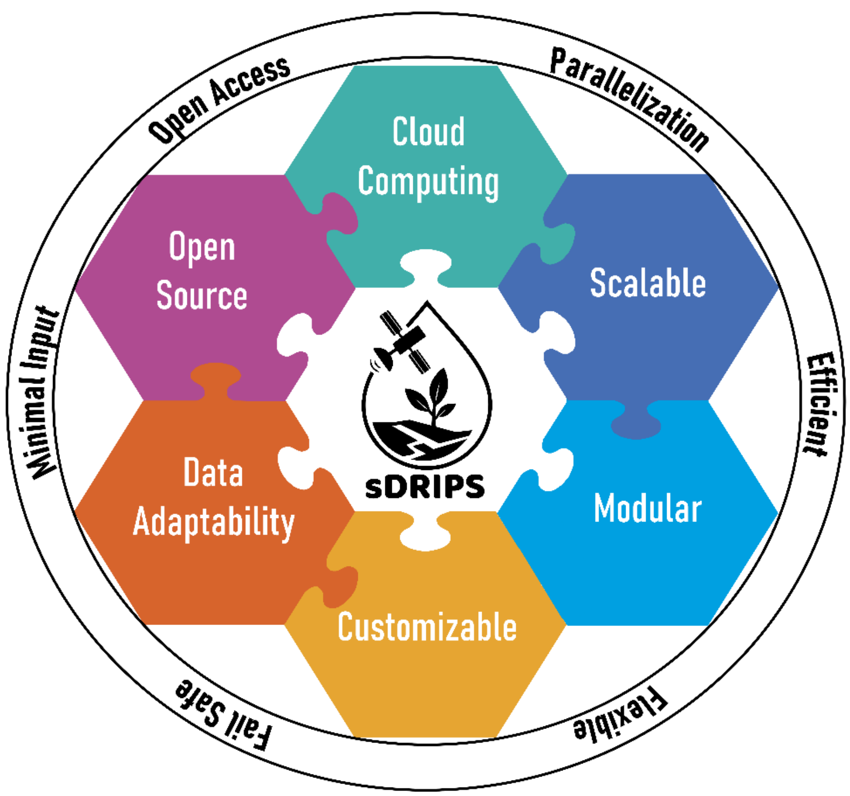

Figure 5.1: Key software design advancements in sDRIPS.

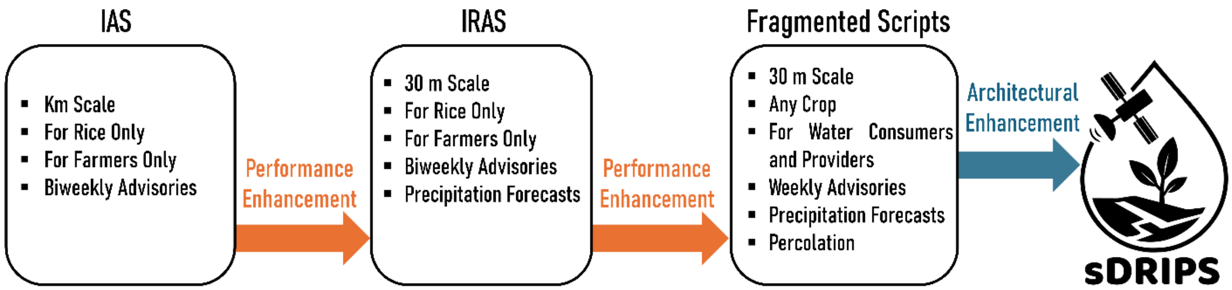

Figure 5.2: Evolution of irrigation advisory systems leading to the development of sDRIPS. The concept originated with the IAS in 2017, which was subsequently expanded to Bangladesh with enhanced spatial resolution (IRAS). Over time, the framework was adapted to multiple regions, operating at a weekly temporal scale, and further evolved into sDRIPS with tailored versions (e.g., sDRIPS-TBP, sDRIPS-Sense, sDRIPS-Ukulima) before being formalized into a unified Python package.

5.5. Building a Python Library

5.5.1. Dependencies

sDRIPS is developed entirely in Python, with its functionalities tightly integrated into the GEE platform via its Python API. This ensures seamless access to cloud-based geospatial datasets while retaining Python's flexibility for scientific computing and data analysis. Importantly, sDRIPS is operating system independent and can be executed across Windows, Linux, and macOS environments.

To ensure smooth functionality, sDRIPS requires a set of core Python dependencies as shown in Table 5.1. These include widely used geospatial and scientific libraries for data handling, numerical computation, and visualization. Dependency management is handled through the standard package manager - pip, which resolves version conflicts and maintains compatibility across different user environments.

Table 5.1: List of Python Dependencies for sDRIPS

Sr. No	Group	Python Dependency	Functionality
1.	Core and Configuration	Python ≥ 3.7	Core interpreter required to execute sDRIPS source code
2.		ruamel.yaml	Reading and writing configuration files in YAML format

3.	Geospatial Data Handling	rasterio	Reading, writing, and processing geospatial raster datasets
4.		rioxarray	Extension of xarray for raster and geospatial operations
5.		geopandas	Handling of vector geospatial data
6.		earthengine-api	GEE Python API
7.		geemap	Using local datasets as GEE assets
8.		cartopy	Cartographic projections and geospatial visualizations
9.	Statistical Tools	numpy	Numerical computations
10.		pandas	Analyzing CSV data
12.		pykrige	Geostatistical interpolation (ordinary and universal kriging)
13.		networkx	For network theory
14.	Utilities	tqdm	Progress bars

5.5.2. System Requirements

A key design principle of sDRIPS is lowering the barrier to entry for users with varying technical expertise and computational resources. By leveraging cloud computing through GEE, the package offloads computationally intensive processes such as satellite image preprocessing and ET estimation using the Penman-Monteith and SEBAL methods to the GEE cloud. Consequently, the computational demand on the user's local machine is modest. On the client side, the package primarily executes water balance calculations (e.g., integrating ET, precipitation, and percolation losses) which do not demand high-performance computing resources. As a result, **sDRIPS can run effectively on a standard laptop** equipped with 4 GB RAM, 512 GB storage, and a dual-core processor. Although more powerful machines equipped with greater RAM and additional CPU cores can reduce runtime, such resources are not strictly necessary for the software's use.

To illustrate the relationship between hardware configuration and runtime performance, we benchmarked sDRIPS using the **Teesta Barrage Project (TBP)** in Bangladesh as a case study. For a single time step, sDRIPS was configured to estimate net water requirements at the command area level, incorporating ET, observed and forecasted precipitation, and percolation losses. Here, a *command area* refers to the land irrigated by a particular canal. Figure 5.3 shows the TBP command area. Table 5.2 summarizes the computational load, while Table 5.3 reports runtime performance under different hardware configurations. To generate Table 5.3 benchmarks, we executed sDRIPS on both Windows and Unix-based platform, testing under minimum-core and maximum-core settings (cores minus one, reserved for system processes). This experimental design allowed us to evaluate performance trade-offs across operating systems and processor configurations, providing guidance for users with varying computational capacities.

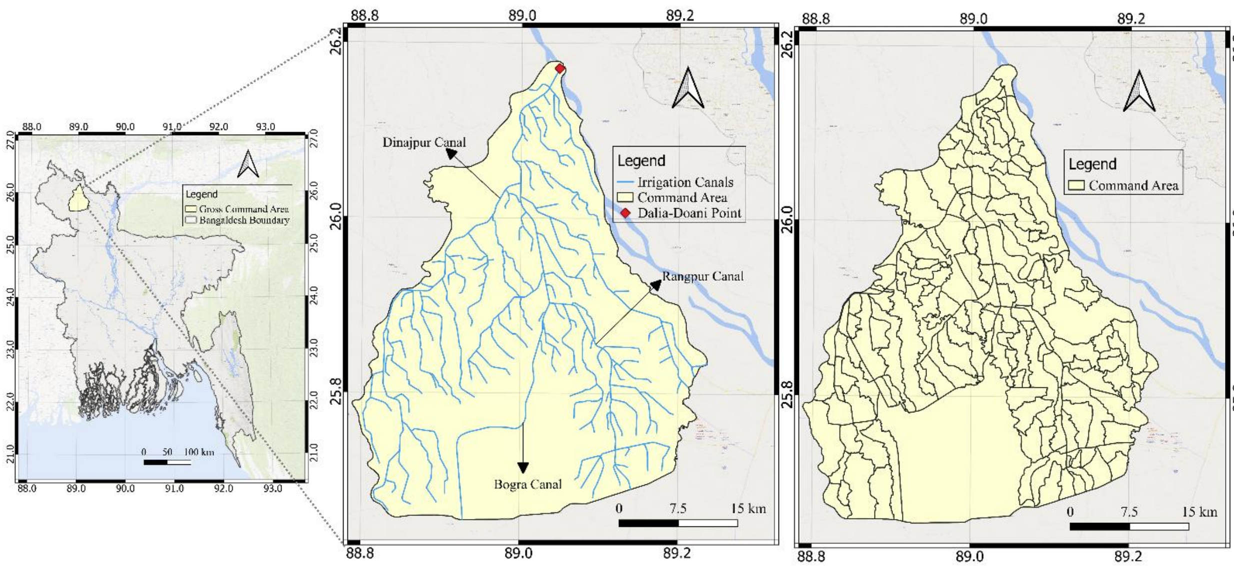

Figure 5.3: Location of TBP in Bangladesh with its canal network and command *areas.*

Table 5.2: Data volume processed for TBP

Sr. No.	Parameter	Value
1.	Number of Command Areas	153
2.	Total Area for Analysis	1548.65 Km^2
3.	Number of Landsat (30m) and Sentinel 1 (10m) Pixels	1720721 and 15486485
4.	Date Selected for Run	28^{th} Feb, 2023

Table 5.3: Computational time comparison of sDRIPS for different machines using the maximum and minimum number of CPU cores.

System Specifications	Windows Laptop		Cluster Machine (UNIX)	
Number of Cores Used	**1**	**Max (8-1)**	**1**	**Max (64-1)**
Evapotranspiration estimation (Penman-Monteith and SEBAL)	49.2 mins	6.3 mins	35.8 mins	4.6 mins
Raster preparation (unzipping and renaming)	2.5 mins	32.4 sec	5.9 sec	5.8 sec
Run time to Precipitation data acquisition and processing (nowcast and forecast)	25.5 sec	12.9 second	5.5 sec	1.7 sec
Percolation estimation	11.2 mins	1.5 mins	9.8 mins	37.6 sec
Net water requirement estimation	3.2 mins	1.9 mins	35.5 sec	37.3 sec
Total run time	**66.6 mins**	**10.4 mins**	**46.4 mins**	**6 mins**

The total runtime in Table 5.3 illustrates the overall benefit of parallelization, the complete workflow required 66.6 minutes on a single-core laptop but only 10.4 minutes when run with parallel processing, and dropped from 46.4 minutes to 6 minutes on the cluster. These benchmarks clearly highlight that sDRIPS can be efficiently deployed across diverse computing environments, making it suitable for both resource-limited users and high-performance operational applications.

5.5.3. Design & Architecture

5.5.3.1. Data Inputs and Prerequisites

The design of sDRIPS emphasizes minimal user-side data requirements while leveraging cloud-hosted datasets for scalability and global applicability. To execute sDRIPS for a given region of interest (ROI), the primary user-provided input is a boundary polygon file (e.g., shapefile or GeoJSON). In addition, users specify the operational mode of sDRIPS, which determines the type of advisories generated:

1. **Field-level advisories for farmers –** Estimation of net water requirements at the plot scale.
2. **Canal-level advisories for water managers** – Estimation of aggregate water demand across a command area, and the amount of water to be supplied to the primary, secondary, and tertiary canals.
3. **Sensor-integrated advisories** – Incorporating in situ soil moisture, air temperature, windspeed, or local weather station observations.
4. **Model-driven advisories** – Using globally available weather and precipitation products when local data are unavailable.

By design, sDRIPS minimizes user data burdens by automatically integrating remote-sensing datasets from Google Earth Engine (such as Landsat series, and Sentinel 1), precipitation estimates from NASA's IMERG, and meteorological inputs from the GFS. When selected, in situ or local weather station data can supplement or replace global model inputs. Thus, sDRIPS adapts seamlessly to varying data availability and user capacity.

Table 5.4 summarizes the minimal data and account credentials required to run sDRIPS, while additional data needs (such as in-situ sensor observations) arise only when enhanced configurations are enabled. For details on the specific type of satellite and model data that goes into sDRIPS, readers are advised to refer to the ReadTheDocs page.

Table 5.4: Table showing minimal data and account (credentials) requirements for running sDRIPS

Sr. No	Requirement	Purpose
1.	Google Earth Engine Account	To utilize the cloud computing power
2.	Precipitation Processing System (PPS) Account	To download the IMERG precipitation dataset
3.	Boundary Polygon (Shapefile/GeoJSON)	Defines the spatial extent of analysis
4.	Crop Type and Planting Date	To accurately estimate the ET
5.	Sensor or Weather Station Data	To accurately estimate the net water requirement. Required only if the user wants to consider the in situ data in the analysis.

5.5.3.2. Processing Pipeline

The centralized architecture of sDRIPS integrates multiple functionalities into a single, modular pipeline. This allows the same framework to serve heterogeneous users – whether water consumers or water providers through configuration-driven execution. A schematic flowchart is shown in Figure 5.4 illustrates the decision logic, where execution branches depend on user-defined configuration settings. At a high level, the processing pipeline consists of five sequential components:

1. **Configuration setup** (orange patch in Figure 5.4) – Reading and validating user-defined parameters from the configuration files.
2. **Evapotranspiration estimation (**purple patch in Figure 5.4**)** – Computing crop water demand using Penman-Monteith and SEBAL models.
3. **Percolation loss estimation** (pink patch in Figure 5.4) – Quantifying percolation losses specific to field conditions.
4. **Precipitation integration (**blue patch in Figure 5.4**) – I**ncorporating both nowcasted and forecasted rainfall from global models or local sources (only nowcast).
5. **Advisory generation** (green patch in Figure 5.4) – Comparing water balance components to produce net water requirement advisories.

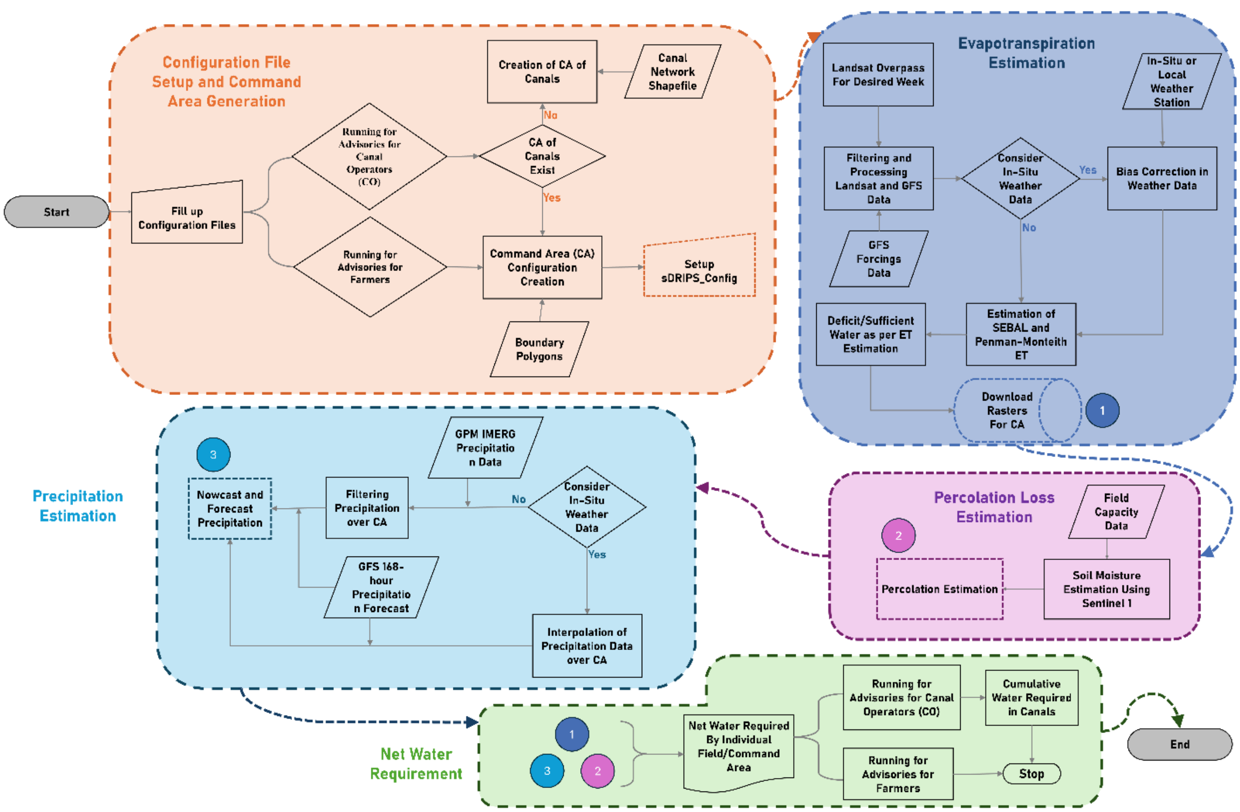

Figure 5.4: Flowchart of the sDRIPS pipeline with the decision logic designed through user-defined configurations.

Each component is modular, enabling users to toggle or customize specific elements without modifying the core workflow. For example, a user may opt to substitute local station rainfall data with IMERG precipitation or disable percolation estimation if not relevant to their context. The detailed computational model, including input specifications and parameterization, is documented on the **ReadTheDocs** page of the package - https://sdrips.readthedocs.io/.

5.5.3.3. Configuration & Modularity

To balance usability with flexibility, sDRIPS is governed by a modular configuration system. The framework is supported by five configuration files, each serving a distinct purpose. Of these, only two require direct user input, while the remaining files store static or system-level information necessary for execution. This modular structure enables users—whether farmers, irrigation district managers, or researchers—to tailor sDRIPS to their specific needs. Importantly, the modular structure also ensures reproducibility and scalability, as configurations can be easily shared or version-controlled. Advanced users may additionally employ the developer version of sDRIPS to introduce new functionalities and contribute to its evolution.

5.5.3.3.1. sDRIPS Configuration File

sDRIPS configuration file (sdrips_config.yaml) provides the primary control interface for running sDRIPS. It includes:

1. **High-Level module controls –** Define the functional mode of sDRIPS (such as generating advisories for farmers versus canal operators). Furthermore, to specify whether to integrate in situ sensor data, local weather station inputs, or rely solely on satellite-based and numerical weather model estimates without bias correction. Overall, High-Level controls govern the core objectives of a model run, and are a configuration switch (‘True’ or ‘False’) that activates a distinct analytical capability.

2. **Low-Level module controls –** Specify common parameters such as the running period, boundary polygon location, project directory paths, directory-cleaning options, etc.

Together, these controls define the scope, inputs, and computational logic of a given sDRIPS run.

5.5.3.3.2. Crop Configuration File

Crop configuration contains crop-specific growth stage information, including crop coefficients (Kc) parameterized by days since planting. Default values for major global crops are sourced from established literature. Users can extend or modify this file to accommodate local varieties or new cultivars with unique growth cycles (such as shorter duration or high-yielding-varieties). This flexibility allows sDRIPS to remain relevant across a wide range of agroecological settings.

5.5.3.3.3. Command Area Configuration File

The command area configuration file defines region-specific management and cropping conditions. It records crop type, planting date, and water distribution efficiency within the field boundary. This enables sDRIPS to model multiple fields or command areas in parallel, even when they differ in cropping patterns or water delivery characteristics.

5.5.3.3.4. Secrets File

Sensitive credentials (PPS credentials or GEE service account information) are stored in a dedicated secrets file. Isolating this information enhances both security and portability, ensuring that sensitive data are not hard-coded into project workflows.

5.5.3.3.5. Link Configurations File

This file stores and updates links to external data services, such as IMERG precipitation datasets. Maintaining these links in a configuration file simplifies updates and ensures that the pipeline remains compatible with evolving data access protocols.

5.6. Enhancements incorporated in the Python library

To broaden accessibility and expand the user base, **sDRIPS has been developed as a Python library** with dual modes of interaction – a command-line interface (CLI) and a Jupyter Notebook environment. This dual design ensures inclusivity across diverse user communities – from farmers and practitioners who may prefer the CLI, to researchers who benefit from the exploratory flexibility of notebooks. The modules follow **PEP 257** docstring conventions for improved readability and include comprehensive help functions detailing input syntax and usage examples wherever possible. To streamline adoption and ensure correct installation, several key functionalities mentioned in the following sections have been incorporated.

5.6.1. Initialization

The setup process of sDRIPS is simplified through the 'sdrips init' command, which automatically downloads the latest configuration files required for execution and pre-fills key parameters. This reduces the manual burden on users. The command requires a path to the **project directory**, where all output data will be stored. Initialization is a **one-time requirement per project**, ensuring that a defined directory structure is followed and sDRIPS is correctly configured.

5.6.2. Testing

Following installation and initialization, users can validate the sDRIPS installation via the ‘sdrips test’ command in the CLI. The test requires IMERG credentials, which are provided in ‘secrets.yaml’, enabling sDRIPS to download IMERG precipitation data. Testing can be performed in two benchmark scenarios:

1. Using satellite and global model data - For the Rupasi region in Bangladesh over a specific time period
2. In situ sensor data integration - University of Washington landscape over a specific time period.

This functionality ensures that sDRIPS has been correctly installed and is fully functional in different computational modes. Beyond user validation, these tests serve a dual role by enabling developers to detect potential pipeline breaks when introducing new functionalities. It also ensures that user credentials in the secrets file are correctly configured.

5.6.3. Configure

While initialization sets up the project directory structure and verifies computational resources (e.g., available CPU cores), certain parameters are unique to each farm or command area – such as crop type, planting date, water distribution efficiency, and region-specific identifiers and cannot be predefined. To accommodate this heterogeneity, sDRIPS gives user the flexibility to create their own command area configuration files automatically, by providing the boundary polygon (a shapefile or a GeoJSON). This ensures that all configuration files complement each other to fully parameterize the pipeline, enabling users to tailor sDRIPS to their specific operational and environmental context.

5.6.4. Execution

Running sDRIPS is straightforward: users execute the ‘sdrips run’ command and specify the path to the configuration file. The library also supports **operational automation,** allowing users to schedule weekly runs via cron jobs on UNIX systems or the Windows Task Scheduler. This flexibility ensures that sDRIPS can be integrated seamlessly into routine workflows.

5.7. Maintenance of the Python library

5.7.1. Documentation

Maintaining the sustainability and usability of **sDRIPS** is critical for long-term adoption and community engagement. Comprehensive documentation has been developed and is hosted on **ReadTheDocs (**https://sdrips.readthedocs.io/), providing a centralized resource to lower the barrier to entry for new users and contributors. The documentation includes the following components:

5.7.1.1. Model Architecture

A detailed overview of the sDRIPS model architecture is provided in two sections – the **Conceptual Model** and the **Computational Model**. Users can refer to these sections to grasp the overall workflow and data processing logic of sDRIPS.

5.7.1.2. User Guide

The user guide offers a comprehensive walkthrough of prerequisites and setup procedures, including account creation for PPS or GEE, complemented by screenshots. It also provides step-by-step instructions for installation, testing, and execution. Furthermore, the user guide also includes points to remember and tips to facilitate user adoption.

5.7.1.3. Guided Tutorials

Guided tutorials demonstrate the usage of sDRIPS through both the CLI and Jupyter Notebook interfaces. These tutorials walk users through each module and provide example setups, enabling new users to configure sDRIPS for their own regions of interest by replicating default case studies.

5.7.1.4. Configuration Parameters

Since configuration files define how sDRIPS operates, the documentation provides a detailed description of each parameter, including its function and default value. It also includes instructions for extending the system, such as adding new crops to the configuration files or updating data links, accompanied by practical examples. Add high and low level controls

5.7.1.5. Developer's Guide

A detailed guide on how to setup the developer version of sDRIPS for contributors who wish to extend or enhance sDRIPS.

5.8. Application of the Python library

This section illustrates the versatility of the sDRIPS Python package across contrasting agroecological and socio-economic contexts. By presenting two distinct use cases, we highlight how sDRIPS can adapt to the availability of i) on-ground resources, ii) scale of operation, and iii) stakeholder needs. The first use case focuses on a developed-world setting where sensor-rich data streams complement the satellites and model products, while the second use case demonstrates its relevance in smallholder farming systems in South Asia. Together, these examples showcase the flexibility of sDRIPS in addressing both water optimization in urban landscapes and irrigation scheduling in resource-constrained agricultural systems. More details are provided in the respective cases.

5.8.1. User Case 1 – sDRIPS-Sense

5.8.1.1. Study Area and Background

This case study was conducted on the lawns of the University of Washington (UW), Seattle, USA. Seattle is located in the Pacific Northwest of the USA, and experiences dry summers and wet winters. Maintaining campus greenery during the dry season often results in excessive irrigation, underscoring the need for data-driven strategies to minimize water waste. Here, the goal of the study was to minimize the water wastage (provide water based on the net water requirement of the grass).

Unlike large, homogeneous agricultural fields, urban landscape patches are small, heterogeneous, and interspersed with built infrastructure such as buildings, roads, and sidewalks. This setting presents two major challenges: i) satellite-based ET estimates are subject to noise from surrounding built-up, and ii) global meteorological reanalysis products often fails to capture the microclimates due to their coarse resolution.

To address these limitations, we deployed in situ sensors measuring air temperature and soil moisture across representative sites within the campus. These ground observations were integrated into sDRIPS, which allows for bias correction of satellite-derived variables and meteorological inputs, allowing sDRIPS to produce reliable ET estimates and net irrigation requirements at the landscape scale.

The case study was implemented during the summer months when irrigation demand is at its peak. Beyond the operational gains for campus facilities management, this case study played a pivotal role in the development of the module enabling seamless integration of in situ observations with satellite-based workflows. Figure 5.5 shows the bounds of the university of Washington.

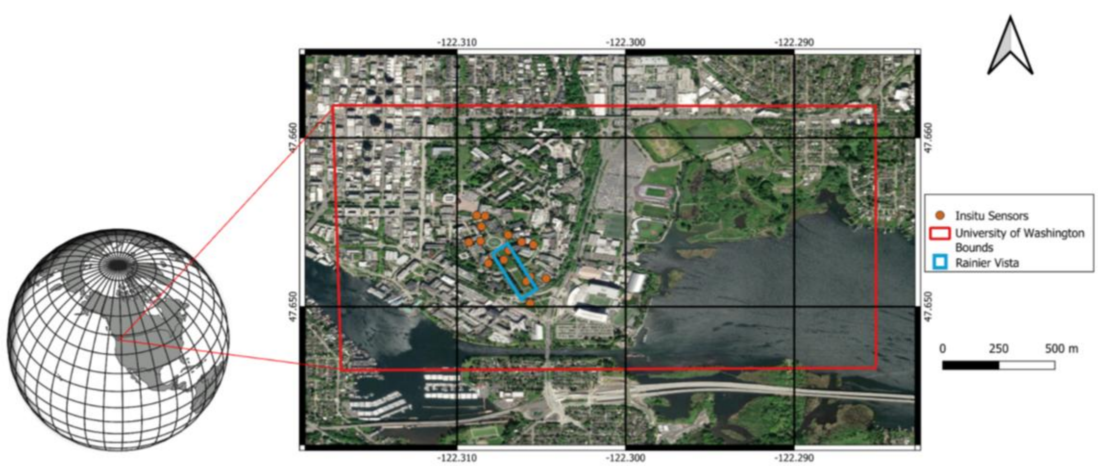

Figure 5.5: University of Washington Campus. The blue box is a region known as Rainier Vista that was used as a test case later.

5.8.1.2. Setup

sDRIPS was deployed over the UW campus for the summer months (May–August). A boundary polygon delineating landscaped areas was derived from satellite imagery to define the study domain. Ground observations were obtained from OnsetComp sensors (https://www.onsetcomp.com/) which measured variables such as air temperature and soil moisture.

The system was configured to operate in a fail-safe hybrid mode. Specifically, when in situ observations were available, sDRIPS integrated them with satellite-derived and global numerical weather model data (GFS) to perform bias correction and improve ET estimates. In cases when sensors failed or data gaps occurred, the system automatically switched to a satellite-only mode without user intervention. To support sensor integration setup, a CSV file containing sensor observations (downloaded from the OnsetComp data hub), and a CSV file specifying sensor locations were provided. These inputs were linked through the sDRIPS configuration file, which maps user-defined variable names (e.g., "Air temperature" or "Temperature" in the CSV) to standardized meteorological variables used by the system. High-Level and Low-Level controls in the configuration file govern the functional mode and execution environment, respectively. For a better understanding of the readers, an excerpt from the sdrips_config.yaml file is provided showing the in-situ sensor configuration section of the sDRIPS. Other configuration settings are provided in the Appendix B.

```
# ------------------------------
# IN-SITU SENSOR CONFIGURATION
# ------------------------------
Insitu_Sensor_integration_config:
  air_temperature_sensor: true            # Use air temperature sensor data for correction. Eg - true or false
  air_temperature_variable: 'Temperature' # Reference name of the variable in insitu data. (If air_temperature_sensor is True)
  soil_moisture_sensor:  false            # Use soil moisture sensor data. Eg - true or false
  soil_moisture_variable: 'Water Content' # Reference name of the variable in insitu data. (If soil_moisture_sensor is True)
  wind_speed_sensor:  true               # Use wind speed sensor data. Eg - true or false
  wind_speed_variable: 'Wind'       # Reference name of the variable in insitu data. (If wind_speed_sensor is True)
  specific_humidity_sensor:  true           # Use specific humidity sensor data. Eg - true or false
  specific_humidity_variable: 'Specific Humidity'       # Reference name of the variable in insitu data. (If specific_humidity_sensor is True)
  hub_url: "https://depts.washington.edu/saswe/Sensor_Cronjob/"  # URL to the sensor data hub (if applicable). Eg - "https://your_data_hub.com/Sensor_Cronjob/"
  date_regex: "(\\d{4}_\\d{2}_\\d{2})" # Regular expression to extract date from sensor filenames. 4 represents _ _ _ _ in Year, 2 represens _ _ in months and another 2 represents _ _ in date.
  date_format: "%Y_%m_%d" # Date format for sensor data filenames. Eg - "%Y_%m_%d" for 2024_04_06
  sensor_data_path:  Sensor_Data                    # Path to folder (should be in save_data_loc folder) containing sensor CSV files
```

```
interpolation_method: best             # Options: 'best', 'average', 'idw', 'kriging',
'inverse_distance_weighting', 'nearest_neighbor'
```

5.8.1.3. Outputs

The primary outputs for this use case is the raster data depicting irrigation requirements across the landscaped areas of the campus. These maps allowed facilities staff to identify patches of under- or over-irrigation and adjust watering schedules accordingly. To facilitate decision-making, outputs were disseminated via a website (https://depts.washington.edu/saswe/sdrips-sense/), and a mobile application (sDRIPS Mobile) for iOS devices (https://apps.apple.com/us/app/uw-waterwise/id6745153515).

The University of Washington case represents a unique setting, as the summer months are characterized by persistent non–rain-bearing cloud cover. To address this, sDRIPS outputs were configured using approximately 20 years of climatological data (estimated through sDRIPS) for the region. Figure 5.6 shows the raster data over the Ranier Vista area (highlighted in blue in Figure 5.5), which was selected due to its high density of in situ sensors. While the present use case highlights this site as an example, landscape workers routinely use the mobile application in summer months to monitor and manage irrigation needs across the broader campus.

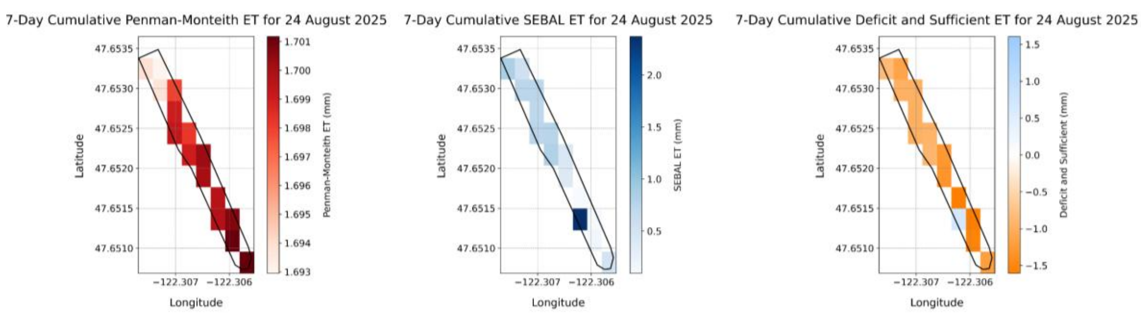

Figure 5.6: ET estimations over the Rainier Vista region (see Figure 5.5).

5.8.2. User Case 2 – TBP

5.8.2.1. Study Area

The second use case illustrates the application of sDRIPS in a resource-constrained agricultural setting without access to in situ sensor networks. The study area corresponds to the command areas in the northern part of Bangladesh supplied by the Teesta Barrage irrigation network (see Figure 5.3). In contrast to the UW case, farmers in this region typically lack access to in situ sensors and rely instead on experiential knowledge to make irrigation decisions. Such subjective practices often result in over- or under-watering, particularly under the increasing variability associated with climate change. Farmers also have limited capacity to anticipate precipitation events, further compounding the challenge.

For this case, the Bangladesh Water Development Board (BWDB) provided shapefiles of the command areas, each encompassing multiple farm fields irrigated via the canal network. While the analysis domain corresponds to that shown in Figure 5.3, the current study focuses on farmer-centric advisories. The advisories then can be disseminated using radio broadcasts and agricultural extension services, which have been successfully used in Bangladesh for similar purposes in the past (*Satellite Data Helps Bangladeshi Farmers Save Water, Money, Energy | Landsat Science*, 2023).

5.8.2.2. Configuration Files

The modular structure of sDRIPS allows adaptation to both farmer-centric and operator-centric applications through minimal configuration changes. For farmer advisories, the sdrips_config.yaml file was set to enable the Command_Area_Net_Water_Requirement module while keeping other components disabled. This setup generated irrigation advisories tailored to guide farmers' in the region. As in User Case 1, an excerpt from the primary configuration file is provided below. Additional technical details regarding Low-Level controls and customization are documented in the Appendix B.

```
# ------------------------------------------
# CONFIGURATION FILE FOR sD.R.I.P.S. (Satellite Data Rendered Irrigation using Penman-Monteith and SEBAL)
# ------------------------------------------


# ------------------------------
# HIGH-LEVEL MODULE CONTROLS
# ------------------------------
Command_Area_Net_Water_Requirement: true  # Generate advisory on how much water should be allotted to each command area
Canal_water_allotment: false          # Generate advisory on how much water should be allotted to each canal (From primary to tertiary canals)
Insitu_Sensor_integration: false  # Toggle to integrate in-situ sensor data with satellite data for bias correction
Weather_station_integration: false  # Toggle to integrate local weather station data to improve accuracy
```

5.8.2.3. Outputs

The outputs consisted of raster data depicting irrigation requirements across all farm fields within the command areas, alongside the CSV file summarizing the irrigation demand. For the configuration used in this run, median irrigation requirements were computed at the command area scale, aggregating across individual farm fields. Nonetheless, the framework retains the flexibility to generate outputs at the farm-field scale, provided that appropriately granular polygons are supplied by the user. Figure 5.7 illustrates the 7-day cumulative Penman-Monteith ET, SEBAL ET, and spatial patterns of deficit and surplus irrigation across a particular command area. For the analysis, ET estimates were generated for the rice crop, assuming a uniform planting date of 15 January 2023 across all command areas. Although sDRIPS is capable of analyzing multiple crops across different farm fields, data limitations on cropping patterns necessitated the assumption of rice for this case study.

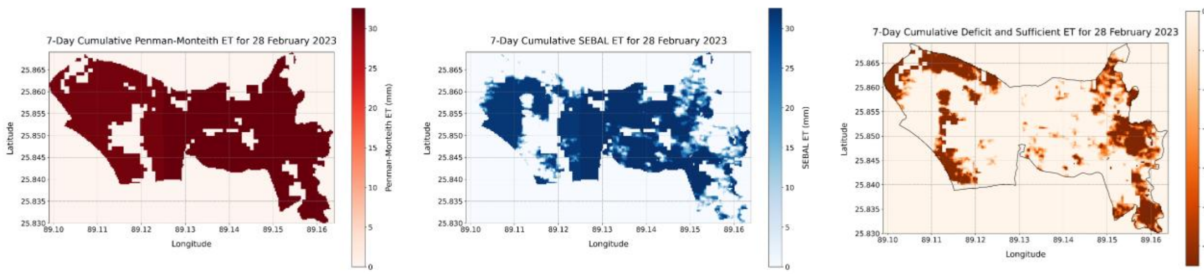

Figure 5.7: sDRIPS raster outputs for a command area in the TBP region.

While sDRIPS also retrieved both nowcast and forecast precipitation datasets during the study period, but no precipitation events occurred, and therefore precipitation rasters are not shown here. In addition to the ET and precipitation, sDRIPS produces further outputs, including percolation losses and net water requirements, which are exported in CSV format. The deficit and surplus maps are particularly valuable as direct decision-support products, allowing farmers to adjust irrigation applications spatially.

5.9. Conclusions and Discussion

Global surface water resources are under increasing stress due to population growth, rising food demand, and climate change. Higher temperatures and altered precipitation regimes are already contributing to irrigation inefficiencies, where both over- and under-watering reduce crop yields while wasting scarce water resources. In this context, scalable, adaptable, and data-driven tools for irrigation management are urgently needed.

This study introduced sDRIPS (satellite Data Rendered Irrigation using Penman-Monteith and SEBAL), a modular, open-source Python package designed to support both water users (e.g., farmers) and water providers (e.g., canal operators). By integrating publicly available satellite observations (e.g., Landsat, Sentinel1, and IMERG), global numerical weather model products (GFS), and optionally in situ or local weather station-based observations, sDRIPS offers a flexible platform for irrigation scheduling across diverse contexts. The centralized design integrates earlier prototype versions into a single unified framework, enabling users with varying technical expertise, operational scales, and data availability to run the same tool in different agricultural settings. To lower barriers to adoption, sDRIPS is supported by detailed documentation and tutorials on ReadTheDocs, while its open-source nature encourages community-driven contributions and innovation.

The use cases demonstrate the adaptability of sDRIPS across contrasting agro-economic settings. In a sensor-rich environment such as the UW campus, sDRIPS optimized the irrigation needs across heterogeneous landscapes by integrating ground-based sensor data with satellite observations. In contrast, within the TBP in northern Bangladesh, a resource-constrained setting where farmers often lack access to sensors. sDRIPS demonstrated the ability to generate actionable advisories at both farm and command-area scales. These advisories can be disseminated via extension channels already familiar to local farmers. Together, these cases underscore sDRIPS's scalability, demonstrating its utility in both high-tech agricultural landscapes and smallholder-dominated irrigation systems.

Beyond operational applications, sDRIPS can also be used for research purposes. Leveraging historical datasets, the tool allows users to simulate "what-if" scenarios such as evaluating the impacts of future policies under present or projected climate conditions, including the transboundary river basins where upstream supply decisions affect downstream availability. This capacity supports science-based policy evaluation and decision-making at the regional and national levels.

Despite its strengths, several limitations remain. First, sDRIPS currently relies on optical satellite data (e.g., Landsat series) to estimate evapotranspiration, making it vulnerable to cloud contamination. Second, while the modular design is user-friendly, and efforts are made to integrate the in-situ sensor data with ease, integrating in situ sensor data from diverse providers still remains challenging. The absence of standardized data formats and processing pipelines across sensor providers complicates interoperability with satellite and weather model products. Addressing these integration barriers will be essential for advancing scalable hybrid systems that combine remote sensing with in-situ monitoring for irrigation management.

In conclusion, sDRIPS represents a step toward bridging the gap between research and practice in surface water irrigation management. By combining transparency, modularity, and scalability, the package empowers a wide spectrum of end users from individual farmers to regional water managers to make more informed decisions. As climate change and water scarcity intensify, open-source frameworks such as sDRIPS can play a pivotal role in advancing sustainable and equitable irrigation practices at local, regional, and global scales.

Chapter 6. Conclusion and Future Recommendations

6.1. Conclusion

This dissertation advances the understanding, monitoring, and management of surface water resources through the synergistic use of satellite remote sensing, in-situ, citizen science data, and cloud computing. A dual-focus approach (*Source-2-Sinks*) that monitors both *sources* (water bodies) and *sinks* (primarily agricultural fields), providing a holistic framework for sustainable water resource management. To accomplish that in Chapter 2, the volume of water stored in water bodies (*sources*) was estimated by combining satellite observations with citizen science-derived gauge measurements. Citizen science not only enhances data availability but also fosters community engagement and sustainable water practices. Furthermore, the uncertainty associated with the volume of water stored in the *source* is also quantified. Recognizing the impracticality of installing gauges at every water body, Chapter 3 introduced a framework to optimize the number and placement of gauges, ensuring efficient network design while maintaining representative monitoring. From Chapter 4 onwards, the focus shifted to the sinks. In Chapter 4, a framework was developed to guide canal operators in allocating water according to net crop water requirements, optimizing irrigation efficiency. Building on this, Chapter 5 presented a standalone Python package targeting both water providers and water consumers, facilitating sustainable irrigation practices globally.

Overall, Chapter 1 outlined the challenges, growing pressures, and proposed approach; Chapters 2 and 3 focused on sustainable monitoring of water sources; and Chapters 4 and 5 emphasized the development of operational frameworks for sustainable surface water use (in *sinks*) and irrigation management. Key takeaways from each chapter are discussed below.

6.1.1. Chapter 2

An ensemble approach combined satellite data with citizen science derived elevation measurements to estimate surface water volumes. Radar (Sentinel 1) and optical (Landsat) sensors complemented each other for continuous monitoring. A time-averaged uncertainty metric highlighted higher errors in mountainous regions due to shadowing and pixel misclassification. The study not only establishes a benchmark for pre-SWOT-era techniques but also has practical applications for water management. By quantifying uncertainty, planners can account for worst-case scenarios such as minimum lake volumes for urban water supply design, and maximum volumes for flood protection infrastructure. These uncertainty bounds provide guidance for managing surface water storage and flood risk in data-scarce regions. While based on pre-SWOT sensors, the framework also informs future missions like SWOT, which are expected to substantially reduce uncertainties.

6.1.2. Chapter 3

The framework integrates hydro-physical factors such as temperature, precipitation, and terrain roughness to determine the optimal number and placement of gauges, while remaining flexible to include additional determinants such as population density, infrastructure, or streamflow. Two methods - the Robust Curve and Hydrologic Parameter Derived (HyPD) modeled lake area-volume relationships, providing benchmark volumes for regional and global assessments. Probabilistic mapping identified gauge locations to maximize representativeness and return on investment, encouraging citizen engagement. The framework also supports future integration with SWOT data for higher-frequency monitoring of small lakes and wetlands.

6.1.3. Chapter 4

sDRIPS TBP - a water provider-centric irrigation advisory system, integrates satellite data and weather forecasts to optimize water allocation across command areas. It generates actionable advisories based on prior irrigation, precipitation, and water availability, producing water stress maps and scenario-based simulations. Validation against OpenET showed strong agreement. A "what-if" analysis indicated that earlier implementation in the Teesta Barrage Project could have conserved ~475 acre-feet of water in one week.

6.1.4. Chapter 5

A Python package was developed for sDRIPS to enable broad accessibility and ease of use across different regions and user groups. The package is fully documented on ReadTheDocs, promoting transparency, reproducibility, and community-driven development. Case studies demonstrated its versatility in sensor-rich (University of Washington) and resource-constrained (Teesta Barrage Project) environments. The sDRIPS Python package leverages cloud computing and parallel processing, operating through user-friendly configuration files that simplify setup and ensure reproducibility. Furthermore, a dedicated developer version has been created to encourage global collaboration, enabling researchers and practitioners to integrate new concepts and advance the use of sDRIPS for sustainable surface water management.

6.1.5. Bridging the gap between research and application

Throughout this dissertation, stakeholder engagement was an integral component of every chapter, ensuring that the scientific outcomes remain relevant and actionable for real-world water management. Continuous interaction with stakeholders was facilitated through multiple training sessions and feedback loops, allowing the research to be refined in response to practical needs and challenges. In Chapter 2, stakeholders from Bangladesh (Bangladesh Water Development Board, BWDB) and Nepal (Department of Hydrology and Meteorology) were trained to utilize satellite-based methods for estimating water surface area. Chapter 3 extended these efforts by updating existing platforms for estimating total surface water volume in Northeastern Bangladesh (hosted here - https://depts.washington.edu/saswe/haors) and developing new platforms for Northwestern Bangladesh (hosted here - https://depts.washington.edu/saswe/barind/) using the Area–Volume curves established in the chapter. In Chapter 4, the framework for the Teesta Barrage Project was operationalized and is currently under consideration for use by BWDB (hosted here - http://depts.washington.edu/saswe/sdrips_tbp/). Chapter 5 incorporated continuous feedback from a private farm (Crown Farms), leading to the implementation of sDRIPS and the creation of a web-based tool (static mirror here - https://depts.washington.edu/saswe/sdrips_ukulima/). A similar platform was developed for the landscape managers at the University of Washington (hosted here - https://uwirrigation.org/), both were accompanied by demonstration sessions/training for effective use.

6.2. Future Recommendations

Based on the findings and contributions of this dissertation, several avenues for future research and practical implementation are recommended to enhance the monitoring and management of our surface water resources:

1. **Quantifying Volumetric Uncertainty through LOCSS Gauges and SWOT satellite:** While this research established pre-SWOT benchmarks for estimating surface water volume and its associated uncertainty using multi-satellite and citizen science observations, the launch of SWOT provides an unprecedented opportunity to refine these analyses. Future work can systematically evaluate the accuracy of SWOT-derived water surface elevations and volumes using the LOCSS gauge network. This will not only validate SWOT's performance across diverse geomorphological settings but also improve uncertainty quantification in regions where environmental complexity, such as mountainous terrain, was shown to drive errors with pre-SWOT sensors. Such integration will close a key observational gap identified in this work (Chapter 2) and move toward globally consistent, uncertainty-aware monitoring of inland surface water bodies.

2. **Fusing Multi-Sensor Observations for Improved Water Classification and Bathymetry Mapping:** The results from this research (Chapter 2) revealed that classification uncertainty remains a limiting factor for volume estimation, and there are classification issues with SWOT as well. To address this, future research should focus on synergistic use of SWOT and high-resolution multispectral data (e.g., Sentinel-2) to improve water pixel classification, especially in challenging environments such as vegetated wetlands or turbid lakes. Additionally based on the learning of Chapter 3, combining SWOT's elevation measurements with long-term satellite altimetry can generate globally consistent bathymetric datasets, enabling derivation of area–

elevation–volume relationships for lakes and reservoirs worldwide. These efforts could possibly fill a crucial gap between surface water extent mapping and volume dynamics, leading to improved temporal continuity and accuracy in hydrologic modeling.

3. **Development of a Crop Atlas Considering Import Costs:** Building on the irrigation decision-support framework developed in this dissertation (Chapter 4 and Chapter 5), future work could focus on creating a global crop-water atlas that integrates sDRIPS outputs with economic and climatic data. This atlas would guide farmers and stakeholders in crop selection by considering historical water demand, economic value, and potential import costs. Such a tool could optimize water use, enhance food security, and support climate-resilient agricultural planning by advising which crops are best suited to local conditions under changing environmental and economic scenarios.

This dissertation represents a journey that spans continents and ideas beginning in the floodplains and wetlands of South Asia, extending to the mountainous lakes of the Pyrenees in France, and ultimately returning to the landscapes of the University of Washington. What began as an effort to monitor and manage water resources across continents ultimately came full circle, bringing these ideas home to the University of Washington. Guided by the Source-to-Sink (S2S) paradigm, this dissertation demonstrates that scientific innovation, when anchored in accessibility, collaboration, and open data, can transcend regional and institutional boundaries. The frameworks and methodologies developed through this work not only advance the science of water monitoring but also embody a philosophy of equitable, transparent, and data driven water stewardship. The path forward for sustainable water management will depend on our ability to connect science with society, and research with responsibility, a hopeful future that rests upon the foundation laid by this dissertation.

References

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, *33*(1), 121–131. https://doi.org/10.1002/joc.3413

Acharya, T. D., Lee, D. H., Yang, I. T., & Lee, J. K. (2016). Identification of Water Bodies in a Landsat 8 OLI Image Using a J48 Decision Tree. *Sensors*, *16*(7), Article 7. https://doi.org/10.3390/s16071075

Acharya, T. D., Subedi, A., & Lee, D. H. (2018). Evaluation of Water Indices for Surface Water Extraction in a Landsat 8 Scene of Nepal. *Sensors (Basel, Switzerland)*, *18*(8), 2580. https://doi.org/10.3390/s18082580

Ahmad, S. K., Hossain, F., Eldardiry, H., & Pavelsky, T. M. (2020). A Fusion Approach for Water Area Classification Using Visible, Near Infrared and Synthetic Aperture Radar for South Asian Conditions. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(4), 2471–2480. IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2019.2950705

Ahmad, S. K., Hossain, F., Pavelsky, T., Parkins, G. M., Yelton, S., Rodgers, M., Little, S., Haldar, D., Ghafoor, S., Khan, R. H., Shawn, N. A., Haque, A., & Biswas, R. K. (2020). Understanding Volumetric Water Storage in Monsoonal Wetlands of Northeastern Bangladesh. *Water Resources Research*, *56*(12), e2020WR027989. https://doi.org/10.1029/2020WR027989

Allen, R., Pereira, L., Raes, D., & Smith, M. (1998). FAO Irrigation and drainage paper No. 56. *Rome: Food and Agriculture Organization of the United Nations*, *56*, 26–40.

Allen, Tasumi, M., & Trezza, R. (2007). Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC)—Model. *Journal of Irrigation and Drainage Engineering*, *133*(4), 380–394. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:4(380)

Alsdorf, D. E., & Lettenmaier, D. P. (2003). Tracking Fresh Water from Space. *Science*. https://doi.org/10.1126/science.1089802

Arias, M., Notarnicola, C., Campo-Bescós, M. Á., Arregui, L. M., & Álvarez-Mozos, J. (2023). Evaluation of soil moisture estimation techniques based on Sentinel-1 observations over wheat fields. *Agricultural Water Management*, *287*, 108422. https://doi.org/10.1016/j.agwat.2023.108422

Ascione, A., Cinque, A., Miccadei, E., Villani, F., & Berti, C. (2008). The Plio-Quaternary uplift of the Apennine chain: New data from the analysis of topography and river valleys in Central Italy. *Geomorphology*, *102*(1), 105–118. https://doi.org/10.1016/j.geomorph.2007.07.022

Baldwin, D., Manfreda, S., Keller, K., & Smithwick, E. A. H. (2017). Predicting root zone soil moisture with soil properties and satellite near-surface moisture data across the conterminous United States. *Journal of Hydrology*, *546*, 393–404. https://doi.org/10.1016/j.jhydrol.2017.01.020

Bastiaanssen, W. G. M., Noordman, E. J. M., Pelgrum, H., Davids, G., Thoreson, B. P., & Allen, R. G. (2005). SEBAL Model with Remotely Sensed Data to Improve Water-Resources Management under Actual Field Conditions. *Journal of Irrigation and Drainage Engineering*, *131*(1), 85–93. https://doi.org/10.1061/(ASCE)0733-9437(2005)131:1(85)

Bastiaanssen, W. G. M., Pelgrum, H., Wang, J., Ma, Y., Moreno, J., Roerink, G., & van der Wal, T. (1998). A remote sensing surface energy balance algorithm for land (SEBAL). *Journal of Hydrology - J HYDROL*, *212*, 213–229.

Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., & Wagner, W. (2019). Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and Overcoming Obstacles. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(1), 520–539. IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2018.2858004

Baup, F., Frappart, F., & Maubant, J. (2014). Combining High-Resolution Satellite Images and Altimetry to Estimate the Volume of Small Lakes. *Hydrology and Earth System Sciences*, *18*(5), 2007–2020. https://doi.org/10.5194/hess-18-2007-2014

Bertini, C., Ridolfi, E., Padua, L., Russo, F., Napolitano, F., & Alfonso, L. (2021). An entropy-based approach for the optimization of rain gauge network using satellite and ground-based data. *Hydrology Research*, *52*. https://doi.org/10.2166/nh.2021.113

Bhogapurapu, N., Dey, S., Homayouni, S., Bhattacharya, A., & Rao, Y. S. (2022). Field-scale soil moisture estimation using sentinel-1 GRD SAR data. *Advances in Space Research*, *70*(12), 3845–3858. https://doi.org/10.1016/j.asr.2022.03.019

Biancamaria, S., Lettenmaier, D. P., & Pavelsky, T. M. (2016). The SWOT Mission and Its Capabilities for Land Hydrology. *Surveys in Geophysics*, *37*(2), 307–337. https://doi.org/10.1007/s10712-015-9346-y

Boretti, A., & Rosa, L. (2019). Reassessing the projections of the World Water Development Report. *Npj Clean Water*, *2*(1), 1–6. https://doi.org/10.1038/s41545-019-0039-9

Bose, I., Hossain, F., Eldardiry, H., Ahmad, S., Biswas, N. K., Bhatti, A. Z., Lee, H., Aziz, M., & Kamal Khan, Md. S. (2021). Integrating Gravimetry Data With Thermal Infra-Red Data From Satellites to Improve Efficiency of Operational Irrigation Advisory in South Asia. *Water Resources Research*, *57*(4), e2020WR028654. https://doi.org/10.1029/2020WR028654

Brakenridge, R., & Anderson, E. (2006). MODIS-Based Flood Detection, Mapping and Measurement: The Potential for Operational Hydrological Applications. In J. Marsalek, G. Stancalie, & G. Balint (Eds.), *Transboundary Floods: Reducing Risks Through Flood Management* (pp. 1–12). Springer Netherlands. https://doi.org/10.1007/1-4020-4902-1_1

Bras, R. L., & Rodríguez-Iturbe, I. (1976). Network design for the estimation of areal mean of rainfall events. *Water Resources Research*, *12*(6), 1185–1195. https://doi.org/10.1029/WR012i006p01185

Bretreger, D., Yeo, I.-Y., Hancock, G., & Willgoose, G. (2020). Monitoring irrigation using landsat observations and climate data over regional scales in the Murray-Darling Basin. *Journal of Hydrology*, *590*, 125356. https://doi.org/10.1016/j.jhydrol.2020.125356

Burn, D. H. (1997). Hydrological information for sustainable development. *Hydrological Sciences Journal*, *42*(4), 481–492. https://doi.org/10.1080/02626669709492048

Calzadilla, A., Rehdanz, K., Betts, R., Falloon, P., Wiltshire, A., & Tol, R. S. J. (2013). Climate change impacts on global agriculture. *Climatic Change*, *120*(1), 357–374. https://doi.org/10.1007/s10584-013-0822-4

Chan, S. K., Bindlish, R., O'Neill, P. E., Njoku, E., Jackson, T., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., Yueh, S., Entekhabi, D., Cosh, M. H., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., … Kerr, Y. (2016). Assessment of the SMAP Passive Soil Moisture Product. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4994–5007. IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2016.2561938

Chen, X., Su, Z., Ma, Y., Trigo, I., & Gentine, P. (2021). Remote Sensing of Global Daily Evapotranspiration based on a Surface Energy Balance Method and Reanalysis Data. *Journal of Geophysical Research: Atmospheres*, *126*(16), e2020JD032873. https://doi.org/10.1029/2020JD032873

Choi, H., & Jeong, J. (2019). Speckle Noise Reduction Technique for SAR Images Using Statistical Characteristics of Speckle Noise and Discrete Wavelet Transform. *Remote Sensing*, *11*(10), Article 10. https://doi.org/10.3390/rs11101184

Cooley, S. W., Ryan, J. C., & Smith, L. C. (2021). Human Alteration of Global Surface Water Storage Variability. *Nature*, *591*(7848), Article 7848. https://doi.org/10.1038/s41586-021-03262-3

Courouble, M., Davidson, N., & Dinesen, L. (2021). *Global Wetland Outlook Special Edition 2021*. Secretariat of the Convention on Wetlands 2021.

Crétaux, J.-F., Abarca-del-Río, R., Bergé-Nguyen, M., Arsen, A., Drolon, V., Clos, G., & Maisongrande, P. (2016). Lake Volume Monitoring from Space. *Surveys in Geophysics*, *37*(2), 269–305. https://doi.org/10.1007/s10712-016-9362-6

Crétaux, J.-F., Bergé-Nguyen, M., Calmant, S., Jamangulova, N., Satylkanov, R., Lyard, F., Perosanz, F., Verron, J., Samine Montazem, A., Le Guilcher, G., Leroux, D., Barrie, J., Maisongrande, P., & Bonnefond, P. (2018). Absolute Calibration or Validation of the Altimeters on the Sentinel-3A and the Jason-3 Over Lake Issykkul (kyrgyzstan). *Remote Sensing*, *10*(11), Article 11. https://doi.org/10.3390/rs10111679

Crétaux, J.-F., & Birkett, C. (2006). Lake Studies from Satellite Radar Altimetry. *Comptes Rendus Geoscience*, *338*(14), 1098–1112. https://doi.org/10.1016/j.crte.2006.08.002

Dai, Q., Bray, M., Zhuo, L., Islam, T., & Han, D. (2017). A Scheme for Rain Gauge Network Design Based on Remotely Sensed Rainfall Measurements. *Journal of Hydrometeorology*, *18*(2), 363–379. https://doi.org/10.1175/JHM-D-16-0136.1

Deines, J. M., Kendall, A. D., Crowley, M. A., Rapp, J., Cardille, J. A., & Hyndman, D. W. (2019). Mapping three decades of annual irrigation across the US High Plains Aquifer using Landsat and Google Earth Engine. *Remote Sensing of Environment*, *233*, 111400. https://doi.org/10.1016/j.rse.2019.111400

D'Odorico, P., Chiarelli, D. D., Rosa, L., Bini, A., Zilberman, D., & Rulli, M. C. (2020). The global value of water in agriculture. *Proceedings of the National Academy of Sciences*, *117*(36), 21985–21993. https://doi.org/10.1073/pnas.2005835117

Dong, Q. (2018). Study on the Crop Irrigation Water Requirement Based on Cropwat in Jinghuiqu Irrigation Area. *IOP Conference Series: Materials Science and Engineering*, *394*(2), 022037. https://doi.org/10.1088/1757-899X/394/2/022037

Downing, J., Prairie, Y., Cole, J., Duarte, C., Tranvik, L., Striegl, R., McDowell, W., Kortelainen, P., Caraco, N. F., Melack, J. M., & Middelburg, J. (2006). The Global Abundance and Size Distribution of Lakes, Ponds, and Impoundments. *Limnology and Oceanography*, *51*, 2388–2397. https://doi.org/10.4319/lo.2006.51.5.2388

Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W., & Li, X. (2016). Water Bodies' Mapping from Sentinel-2 Imagery with Modified Normalized Difference Water Index at 10-m Spatial Resolution Produced by Sharpening the SWIR Band. *Remote Sensing*, *8*(4), Article 4. https://doi.org/10.3390/rs8040354

Du, Z., Li, W., Zhou, D., Tian, L., Ling, F., Wang, H., Gui, Y., & Sun, B. (2014). Analysis of Landsat-8 Oli Imagery for Land Surface Water Mapping. *Remote Sensing Letters*, *5*(7), 672–681. https://doi.org/10.1080/2150704X.2014.960606

Duan, Z., & Bastiaanssen, W. G. M. (2013). Estimating Water Volume Variations in Lakes and Reservoirs from Four Operational Satellite Altimetry Databases and Satellite Imagery Data. *Remote Sensing of Environment*, *134*, 403–416. https://doi.org/10.1016/j.rse.2013.03.010

Eagleson, P. S. (1967). Optimum density of rainfall networks. *Water Resources Research*, *3*(4), 1021–1033. https://doi.org/10.1029/WR003i004p01021

Ebert, K. (2015). GIS analyses of ice-sheet erosional impacts on the exposed shield of Baffin Island, eastern Canadian Arctic. *Canadian Journal of Earth Sciences*, 150722143659002. https://doi.org/10.1139/cjes-2015-0063

Farhat, Y. A., Kim, S.-H., Seyfferth, A. L., Zhang, L., & Neumann, R. B. (2021). Altered arsenic availability, uptake, and allocation in rice under elevated temperature. *Science of The Total Environment*, *763*, 143049. https://doi.org/10.1016/j.scitotenv.2020.143049

Feng, W., Zhong, M., Lemoine, J.-M., Biancale, R., Hsu, H.-T., & Xia, J. (2013). Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements. *Water Resources Research*, *49*(4), 2110–2118. https://doi.org/10.1002/wrcr.20192

Fischer, G., Tubiello, F. N., van Velthuizen, H., & Wiberg, D. A. (2007). Climate change impacts on irrigation water requirements: Effects of mitigation, 1990–2080. *Technological Forecasting and Social Change*, *74*(7), 1083–1107. https://doi.org/10.1016/j.techfore.2006.05.021

Forcén-Muñoz, M., Pavón-Pulido, N., López-Riquelme, J. A., Temnani-Rajjaf, A., Berríos, P., Morais, R., & Pérez-Pastor, A. (2022). Irriman Platform: Enhancing Farming Sustainability through Cloud Computing Techniques for Irrigation Management. *Sensors*, *22*(1), Article 1. https://doi.org/10.3390/s22010228

Frazier, P. S., & Page, K. J. (2000). Water Body Detection and Delineation with Landsat TM Data. *Photogrammetric Engineering and Remote Sensing*, *66*, 1461–1467.

Gabr, M. E., & Fattouh, E. M. (2021). Assessment of irrigation management practices using FAO-CROPWAT 8, case studies: Tina Plain and East South El-Kantara, Sinai, Egypt. *Ain Shams Engineering Journal*, *12*(2), 1623–1636. https://doi.org/10.1016/j.asej.2020.09.017

Gleick, P. H. (Ed.). (2014). *The World's Water*. Island Press/Center for Resource Economics. https://doi.org/10.5822/978-1-61091-483-3

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Haque, Md. N., Siddika, S., Sresto, M. A., Saroar, Md. M., & Shabab, K. R. (2021). Geo-spatial Analysis for Flash Flood Susceptibility Mapping in the North-East Haor (Wetland) Region in Bangladesh. *Earth Systems and Environment*, *5*(2), 365–384. https://doi.org/10.1007/s41748-021-00221-w

Hengl, T., & Gupta, S. (2019). *Soil water content (volumetric %) for 33kPa and 1500kPa suctions predicted at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution* (Version v0.1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.2784001

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hossain, F., Biswas, N. K., Ashraf, M., & Bhatti, A. Z. (2017, June 21). *Growing More with Less Using Cell Phones and Satellite Data*. Eos. http://eos.org/science-updates/growing-more-with-less-using-cell-phones-and-satellite-data

Hossain, F., Harsha, S. K., Goyal, S., Ahmad, S., Lohani, B., Balaji, N., & Tripathi, S. (2020). Towards affordable and sustainable water-smart irrigation services. *Water Resources IMPACT*, *Vol 22, January 2020*, pp 28-29. https://scholar.google.com/scholar_lookup?title=Towards%20affordable%20and%20sustaina ble%20water-smart%20irrigation%20services&publication_year=2020&author=F.%20Hossein&author=Sri%20Harsha&author=K.%20Goyal&author=S.%20Ahmad&author=S.K.%20Lohani&author=B.%20Balaji&author=N.%20Tripathi%2C%20S

Hossain, F., Hoque, B. A., Ahmed, T., Khanam, S., Biswas, N. K., Akanda, A. S., Khan, S. K., & Katagami, M. (2022). Impact Evaluation of an Operational Satellite-based Integrated Rice Advisory System in Northeastern Bangladesh. *International Journal of Irrigation and Water Management*, *Vol 9 (1)*, 001–011. https://www.internationalscholarsjournals.org/articles/2241174506012022

Hossain, Md. N., Chowdhury, S., & Paul, S. K. (2016). Farmer-level adaptation to climate change and agricultural drought: Empirical evidences from the Barind region of Bangladesh. *Natural Hazards*, *83*(2), 1007–1026. https://doi.org/10.1007/s11069-016-2360-7

Hu, S., Niu, Z., Chen, Y., Li, L., & Zhang, H. (2017). Global Wetlands: Potential Distribution, Wetland Loss, and Status. *Science of The Total Environment*, *586*, 319–327. https://doi.org/10.1016/j.scitotenv.2017.02.001

Huang, C., Chen, Y., Wu, J., Li, L., & Liu, R. (2015). An Evaluation of Suomi NPP-VIIRS Data for Surface Water Detection. *Remote Sensing Letters*, *6*(2), 155–164. https://doi.org/10.1080/2150704X.2015.1017664

Huang, W., DeVries, B., Huang, C., Lang, M. W., Jones, J. W., Creed, I. F., & Carroll, M. L. (2018). Automated Extraction of Surface Water Extent from Sentinel-1 Data. *Remote Sensing*, *10*(5), Article 5. https://doi.org/10.3390/rs10050797

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., & Xie, P. (2015). Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). In *ResearchGate*. https://doi.org/10.1007/978-3-030-24568-9_19

Hulley, G. C., Göttsche, F. M., Rivera, G., Hook, S. J., Freepartner, R. J., Martin, M. A., Cawse-Nicholson, K., & Johnson, W. R. (2022). Validation and Quality Assessment of the ECOSTRESS Level-2 Land Surface Temperature and Emissivity Product. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–23. https://doi.org/10.1109/TGRS.2021.3079879

Islam, A. R. Md. T., Karim, Md. R., & Mondol, M. A. H. (2021). Appraising trends and forecasting of hydroclimatic variables in the north and northeast regions of Bangladesh. *Theoretical and Applied Climatology*, *143*(1), 33–50. https://doi.org/10.1007/s00704-020-03411-0

Islam, S., & Gnauck, A. (2008). Mangrove Wetland Ecosystems in Ganges-Brahmaputra Delta in Bangladesh. *Front Earth Sci China*, *2*, 439–448. https://doi.org/10.1007/s11707-008-0049-2

Islam, S., & Ma, M. (2018). Geospatial Monitoring of Land Surface Temperature Effects on Vegetation Dynamics in the Southeastern Region of Bangladesh from 2001 to 2016. *ISPRS International Journal of Geo-Information*, *7*(12), Article 12. https://doi.org/10.3390/ijgi7120486

Jasechko, S., Seybold, H., Perrone, D., Fan, Y., Shamsudduha, M., Taylor, R. G., Fallatah, O., & Kirchner, J. W. (2024). Rapid groundwater decline and some cases of recovery in aquifers globally. *Nature*, *625*(7996), 715–721. https://doi.org/10.1038/s41586-023-06879-8

Ji, L., Gong, P., Geng, X., & Zhao, Y. (2015). Improving the Accuracy of the Water Surface Cover Type in the 30 m FROM-GLC Product. *Remote Sensing*, *7*(10), Article 10. https://doi.org/10.3390/rs71013507

Jones, J. W. (2019). Improved Automated Detection of Subpixel-Scale Inundation—Revised Dynamic Surface Water Extent (DSWE) Partial Surface Water Tests. *Remote Sensing*, *11*(4), Article 4. https://doi.org/10.3390/rs11040374

Kamruzzaman, Md., & Shaw, R. (2018). Flood and Sustainable Agriculture in the Haor Basin of Bangladesh: A Review Paper. *Universal Journal of Agricultural Research*, *6*(1), 40–49. https://doi.org/10.13189/ujar.2018.060106

Karimzadeh, S., Miyajima, M., Kamel, B., & Pessina, V. (2015). A fast topographic characterization of seismic station locations in Iran through integrated use of digital elevation models and GIS. *Journal of Seismology*, *19*(4), 949–967. https://doi.org/10.1007/s10950-015-9505-0

Karthikeyan, L., Chawla, I., & Mishra, A. K. (2020). A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. *Journal of Hydrology*, *586*, 124905. https://doi.org/10.1016/j.jhydrol.2020.124905

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S. E., Leroux, D., Mialon, A., & Delwart, S. (2012). The SMOS Soil Moisture Retrieval Algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1384–1403. IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2012.2184548

Keum, J., & Coulibaly, P. (2017). Information theory-based decision support system for integrated design of multivariable hydrometric networks. *Water Resources Research*, *53*(7), 6239–6259. https://doi.org/10.1002/2016WR019981

Khan, S., Hossain, F., Pavelsky, T., Parkins, G., Lane, M., Gomez, A., Minocha, S., Das, P., Ghafoor, S., Bhuyan, M., Haque, M., Sarker, P., Borua, P., Cretaux, J., Picot, N., Balakrishnan, V., Ahmad, S., Thapa, N., Bhattarai, R., & Compin, A. (2023). Understanding Volume Estimation Uncertainty of Lakes and Wetlands Using Satellites and Citizen Science. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *PP*, 1–22. https://doi.org/10.1109/JSTARS.2023.3250354

Khan, S., Hossain, F., Pavelsky, T., Parkins, G. M., Lane, M. R., Gómez, A. M., Minocha, S., Das, P., Ghafoor, S., Bhuyan, Md. A., Haque, Md. N., Sarker, P. K., Borua, P. P., Cretaux, J.-F., Picot, N., Balakrishnan, V., Ahmad, S., Thapa, N., Bhattarai, R., … Compin, A. (2023). Understanding Volume Estimation Uncertainty of Lakes and Wetlands Using Satellites and Citizen Science. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *16*, 2386–2401. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2023.3250354

Khaydar, D., Chen, X., Huang, Y., Ilkhom, M., Liu, T., Friday, O., Farkhod, A., Khusen, G., & Gulkaiyr, O. (2021). Investigation of crop evapotranspiration and irrigation water requirement in the lower Amu Darya River Basin, Central Asia. *Journal of Arid Land*, *13*(1), 23–39. https://doi.org/10.1007/s40333-021-0054-9

Kim, Y., Park, H., Kimball, J. S., Colliander, A., & McCabe, M. F. (2023). Global estimates of daily evapotranspiration using SMAP surface and root-zone soil moisture. *Remote Sensing of Environment*, *298*, 113803. https://doi.org/10.1016/j.rse.2023.113803

Klogo, G., Gasonoo, A., & Ampomah, I. (2013). On the Performance of Filters for Reduction of Speckle Noise in SAR Images Off the Coast of the Gulf of Guinea. *International Journal of*

Information Technology, Modeling and Computing, *1*. https://doi.org/10.5121/ijitmc.2013.1405

Kovenock, M., & Swann, A. L. S. (2018). Leaf Trait Acclimation Amplifies Simulated Climate Warming in Response to Elevated Carbon Dioxide. *Global Biogeochemical Cycles*, *32*(10), 1437–1448. https://doi.org/10.1029/2018GB005883

Krajewski, W. F., Lakshmi, V., Georgakakos, K. P., & Jain, S. C. (1991). A Monte Carlo Study of rainfall sampling effect on a distributed catchment model. *Water Resources Research*, *27*(1), 119–128. https://doi.org/10.1029/90WR01977

Kuang, X., Liu, J., Scanlon, B. R., Jiao, J. J., Jasechko, S., Lancia, M., Biskaborn, B. K., Wada, Y., Li, H., Zeng, Z., Guo, Z., Yao, Y., Gleeson, T., Nicot, J.-P., Luo, X., Zou, Y., & Zheng, C. (2024). The changing nature of groundwater in the global water cycle. *Science*, *383*(6686), eadf0630. https://doi.org/10.1126/science.adf0630

Lehner, B., & Döll, P. (2004). Development and Validation of a Global Database of Lakes, Reservoirs and Wetlands. *Journal of Hydrology*, *296*(1), 1–22. https://doi.org/10.1016/j.jhydrol.2004.03.028

Leonard, P. B., Baldwin, R. F., Homyack, J. A., & Wigley, T. B. (2012). Remote detection of small wetlands in the Atlantic coastal plain of North America: Local relief models, ground validation, and high-throughput computing. *Forest Ecology and Management*, *284*, 107–115. https://doi.org/10.1016/j.foreco.2012.07.034

Lettenmaier, D. P., Alsdorf, D., Dozier, J., Huffman, G. J., Pan, M., & Wood, E. F. (2015). Inroads of Remote Sensing into Hydrologic Science During the Wrr Era. *Water Resources Research*, *51*(9), 7309–7342. https://doi.org/10.1002/2015WR017616

Liang, J., & Liu, D. (2020). A Local Thresholding Approach to Flood Water Delineation Using Sentinel-1 Sar Imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *159*, 53–62. https://doi.org/10.1016/j.isprsjprs.2019.10.017

Little, S., Pavelsky, T. M., Hossain, F., Ghafoor, S., Parkins, G. M., Yelton, S. K., Rodgers, M., Yang, X., Crétaux, J.-F., Hein, C., Ullah, M. A., Lina, D. H., Thiede, H., Kelly, D., Wilson, D., & Topp, S. N. (2021). Monitoring Variations in Lake Water Storage with Satellite Imagery and Citizen Science. *Water*, *13*(7), Article 7. https://doi.org/10.3390/w13070949

Liu, C. (2016). *Analysis of Sentinel-1 Sar Data for Mapping Standing Water in the Twente Region*. 47.

Liu, J., Zhou, Z., Yan, Z., Gong, J., Jia, Y., Xu, C.-Y., & Wang, H. (2019). A new approach to separating the impacts of climate change and multiple human activities on water cycle processes based on a distributed hydrological model. *Journal of Hydrology*, *578*, 124096. https://doi.org/10.1016/j.jhydrol.2019.124096

Liu, P.-W., Famiglietti, J. S., Purdy, A. J., Adams, K. H., McEvoy, A. L., Reager, J. T., Bindlish, R., Wiese, D. N., David, C. H., & Rodell, M. (2022). Groundwater depletion in California's Central Valley accelerates during megadrought. *Nature Communications*, *13*(1), 7825. https://doi.org/10.1038/s41467-022-35582-x

Lowry, C. S., & Fienen, M. N. (2013). Crowdhydrology: Crowdsourcing Hydrologic Data and Engaging Citizen Scientists. *Ground Water*, *51*(1), 151–156. https://doi.org/10.1111/j.1745-6584.2012.00956.x

Mayr, S., Klein, I., Rutzinger, M., & Kuenzer, C. (2021). Determining Temporal Uncertainty of a Global Inland Surface Water Time Series. *Remote Sensing*, *13*(17), Article 17. https://doi.org/10.3390/rs13173454

McCullough, I. M., Loftin, C. S., & Sader, S. A. (2012). High-Frequency Remote Monitoring of Large Lakes with MODIS 500m Imagery. *Remote Sensing of Environment*, *124*, 234–241. https://doi.org/10.1016/j.rse.2012.05.018

McFeeters, S. K. (1996). The Use of the Normalized Difference Water Index (ndwi) in the Delineation of Open Water Features. *International Journal of Remote Sensing*, *17*(7), 1425–1432. https://doi.org/10.1080/01431169608948714

Mekonnen, M. M., & Hoekstra, A. Y. (2016). Four billion people facing severe water scarcity. *Science Advances*, *2*(2), e1500323. https://doi.org/10.1126/sciadv.1500323

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the Volume and Age of Water Stored in Global Lakes Using a Geo-Statistical Approach. *Nature Communications*, *7*(1), Article 1. https://doi.org/10.1038/ncomms13603

Mishra, A. K., & Coulibaly, P. (2009). Developments in hydrometric network design: A review. *Reviews of Geophysics*, *47*(2). https://doi.org/10.1029/2007RG000243

Mishra, A. K., & Coulibaly, P. (2014). Variability in Canadian Seasonal Streamflow Information and Its Implication for Hydrometric Network Design. *Journal of Hydrologic Engineering*, *19*(8), 05014003. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000971

Morsy, M., Taghizadeh-Mehrjardi, R., Michaelides, S., Scholten, T., Dietrich, P., & Schmidt, K. (2021). Optimization of Rain Gauge Networks for Arid Regions Based on Remote Sensing Data. *Remote Sensing*, *13*(21), Article 21. https://doi.org/10.3390/rs13214243

NASA/METI/AIST/Japan Spacesystems And U.S./Japan ASTER Science Team. (2019). *ASTER Global Digital Elevation Model V003* [Dataset]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/ASTER/ASTGTM.003

Padrón, R. S., Gudmundsson, L., Decharme, B., Ducharne, A., Lawrence, D. M., Mao, J., Peano, D., Krinner, G., Kim, H., & Seneviratne, S. I. (2020). Observed changes in dry-season water availability attributed to human-induced climate change. *Nature Geoscience*, *13*(7), 477–481. https://doi.org/10.1038/s41561-020-0594-1

Pareeth, S., & Karimi, P. (2023). Evapotranspiration estimation using Surface Energy Balance Model and medium resolution satellite data: An operational approach for continuous monitoring. *Scientific Reports*, *13*(1), 12026. https://doi.org/10.1038/s41598-023-38563-2

Paul, S., & Hasan, K. (2021). The impact of the proposed rubber dam facilitated surface water irrigation on adjacent groundwater at Chapai Nawabganj district, Bangladesh. *Applied Water Science*, *11*(2), 36. https://doi.org/10.1007/s13201-021-01369-6

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422. https://doi.org/10.1038/nature20584

Pilon, P. J., Yuzyk, T. R., Hale, R. A., & Day, T. J. (1996). Challenges Facing Surface Water Monitoring in Canada. *Canadian Water Resources Journal / Revue Canadienne Des Ressources Hydriques*, *21*(2), 157–164. https://doi.org/10.4296/cwrj2102157

Portoghese, I., Giannoccaro, G., Giordano, R., & Pagano, A. (2021). Modeling the impacts of volumetric water pricing in irrigation districts with conjunctive use of surface and groundwater resources. *Agricultural Water Management*, *244*, 106561. https://doi.org/10.1016/j.agwat.2020.106561

Qayyum, N., Ghuffar, S., Ahmad, H. M., Yousaf, A., & Shahid, I. (2020). Glacial Lakes Mapping Using Multi Satellite PlanetScope Imagery and Deep Learning. *ISPRS International Journal of Geo-Information*, *9*(10), Article 10. https://doi.org/10.3390/ijgi9100560

Qin, J., Duan, W., Zou, S., Chen, Y., Huang, W., & Rosa, L. (2024). Global energy use and carbon emissions from irrigated agriculture. *Nature Communications*, *15*(1), 3084. https://doi.org/10.1038/s41467-024-47383-5

Rahaman, A. Z., Sarker, G. C., Hossain, B. M. T. A., Bhuiyan, S. R., & Naem, A. S. M. J. (2021). Assessment of Community Based Climate Change Risk Focusing Agriculture and Fisheries Sector in Haor Areas of Bangladesh. *Atmospheric and Climate Sciences*, *11*(02), Article 02. https://doi.org/10.4236/acs.2021.112020

Rawat, K. S., Bala, A., Singh, S. K., & Pal, R. K. (2017). Quantification of wheat crop evapotranspiration and mapping: A case study from Bhiwani District of Haryana, India. *Agricultural Water Management*, *187*, 200–209. https://doi.org/10.1016/j.agwat.2017.03.015

Richter, B. D., Lamsal, G., Marston, L., Dhakal, S., Sangha, L. S., Rushforth, R. R., Wei, D., Ruddell, B. L., Davis, K. F., Hernandez-Cruz, A., Sandoval-Solis, S., & Schmidt, J. C. (2024). New water accounting reveals why the Colorado River no longer reaches the sea. *Communications Earth & Environment*, *5*(1), 1–12. https://doi.org/10.1038/s43247-024-01291-0

Riley, J. W., Calhoun, D. L., Barichivich, W. J., & Walls, S. C. (2017). Identifying Small Depressional Wetlands and Using a Topographic Position Index to Infer Hydroperiod Regimes for Pond-Breeding Amphibians. *Wetlands*, *37*(2), 325–338. https://doi.org/10.1007/s13157-016-0872-2

RIVER RESEARCH INSTITUTE. (2023). *River Research Insititute Annual Report 2022-2023* (Annual Report No. 50).

Rodríguez-Iturbe, I., & Mejía, J. M. (1974). The design of rainfall networks in time and space. *Water Resources Research*, *10*(4), 713–728. https://doi.org/10.1029/WR010i004p00713

Ruszkiczay-Rüdiger, Z., Fodor, L., Horváth, E., & Telbisz, T. (2009). Discrimination of fluvial, eolian and neotectonic features in a low hilly landscape: A DEM-based morphotectonic analysis in the Central Pannonian Basin, Hungary. *Geomorphology*, *104*, 203–217. https://doi.org/10.1016/j.geomorph.2008.08.014

Satellite Data Helps Bangladeshi Farmers Save Water, Money, Energy | Landsat Science. (2023, August 3). https://landsat.gsfc.nasa.gov/article/satellite-data-helps-bangladeshi-farmers-save-water-money-energy/

Sawassi, A., Ottomano Palmisano, G., Crookston, B., & Khadra, R. (2022). The Dominance-based Rough Set Approach for analysing patterns of flexibility allocation and design-cost criteria in large-scale irrigation systems. *Agricultural Water Management*, *272*, 107842. https://doi.org/10.1016/j.agwat.2022.107842

Shen, X., Wang, D., Mao, K., Anagnostou, E., & Hong, Y. (2019). Inundation Extent Mapping by Synthetic Aperture Radar: A Review. *Remote Sensing*, *11*(7), Article 7. https://doi.org/10.3390/rs11070879

Sheng, Y., Song, C., Wang, J., Lyons, E. A., Knox, B. R., Cox, J. S., & Gao, F. (2016). Representative Lake Water Extent Mapping at Continental Scales Using Multi-Temporal Landsat-8 Imagery. *Remote Sensing of Environment*, *185*, 129–141. https://doi.org/10.1016/j.rse.2015.12.041

Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of Remote Sensing in Precision Agriculture: A Review. *Remote Sensing*, *12*(19), Article 19. https://doi.org/10.3390/rs12193136

Smarter irrigation for a greener UW | UWCEE. (2024, August 26). *UW Department of Civil & Environmental Engineering*. https://www.ce.washington.edu/news/article/2024-08-27/smarter-irrigation-greener-uw

Smith, M. (1992). *CROPWAT: A Computer Program for Irrigation Planning and Management*. Food & Agriculture Org.

Soares, L. M. V., & Calijuri, M. do C. (2021). Deterministic modelling of freshwater lakes and reservoirs: Current trends and recent progress. *Environmental Modelling & Software*, *144*, 105143. https://doi.org/10.1016/j.envsoft.2021.105143

Solangi, G. S., Shah, S. A., Alharbi, R. S., Panhwar, S., Keerio, H. A., Kim, T.-W., Memon, J. A., & Bughio, A. D. (2022). Investigation of Irrigation Water Requirements for Major Crops Using CROPWAT Model Based on Climate Data. *Water*, *14*(16), Article 16. https://doi.org/10.3390/w14162578

Surendran, U., Sushanth, C. M., Mammen, G., & Joseph, E. J. (2015). Modelling the Crop Water Requirement Using FAO-CROPWAT and Assessment of Water Resources for Sustainable Water Resource Management: A Case Study in Palakkad District of Humid Tropical Kerala, India. *Aquatic Procedia*, *4*, 1211–1219. https://doi.org/10.1016/j.aqpro.2015.02.154

Tagil, S., & Jenness, J. (2008). GIS-Based Automated Landform Classification and Topographic, Landcover and Geologic Attributes of Landforms Around the Yazoren Polje, Turkey. *Journal of Applied Sciences*, *8*, 910–921. https://doi.org/10.3923/jas.2008.910.921

Tang, R., Li, Z.-L., Chen, K.-S., Jia, Y., Li, C., & Sun, X. (2013). Spatial-scale effect on the SEBAL model for evapotranspiration estimation using remote sensing data. *Agricultural and Forest Meteorology*, *174–175*, 28–42. https://doi.org/10.1016/j.agrformet.2013.01.008

Tasumi, M. (2019). Estimating evapotranspiration using METRIC model and Landsat data for better understandings of regional hydrology in the western Urmia Lake Basin. *Agricultural Water Management*, *226*, 105805. https://doi.org/10.1016/j.agwat.2019.105805

Tyralis, H., Papacharalampous, G., & Langousis, A. (2019). A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources. *Water*, *11*(5), Article 5. https://doi.org/10.3390/w11050910

Vermote, E., Roger, J. C., Franch, B., & Skakun, S. (2018). Lasrc (Land Surface Reflectance Code): Overview, Application and Validation Using MODIS, VIIRS, Landsat and Sentinel 2 Data's. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 8173–8176. https://doi.org/10.1109/IGARSS.2018.8517622

Verpoorter, C., Kutser, T., Seekell, D. A., & Tranvik, L. J. (2014). A Global Inventory of Lakes Based on High-Resolution Satellite Imagery. *Geophysical Research Letters*, *41*(18), 6396–6402. https://doi.org/10.1002/2014GL060641

Volk, J. M., Huntington, J. L., Melton, F. S., Allen, R., Anderson, M., Fisher, J. B., Kilic, A., Ruhoff, A., Senay, G. B., Minor, B., Morton, C., Ott, T., Johnson, L., Comini de Andrade, B., Carrara, W., Doherty, C. T., Dunkerly, C., Friedrichs, M., Guzman, A., … Yang, Y. (2024). Assessing the accuracy of OpenET satellite-based evapotranspiration data to support water resource and land management applications. *Nature Water*, *2*(2), 193–205. https://doi.org/10.1038/s44221-023-00181-7

Wagner, W., Lemoine, G., Borgeaud, M., & Rott, H. (1999). A study of vegetation cover effects on ERS scatterometer data. *IEEE Transactions on Geoscience and Remote Sensing*, *37*(2), 938–948. IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/36.752212

Weiss, A. D. (2001). *Topographic position and landforms analysis*. https://www.semanticscholar.org/paper/Topographic-position-and-landforms-analysis-Weiss/221157f21f5f9c479de7328f03f980251708af09

Wine, M. L., & Laronne, J. B. (2020). In Water-Limited Landscapes, an Anthropocene Exchange: Trading Lakes for Irrigated Agriculture. *Earth's Future*, *8*(4), e2019EF001274. https://doi.org/10.1029/2019EF001274

World Meteorological Organization (Ed.). (2008). *Guide to hydrological practices* (6th ed). WMO.

Wu, H., Chen, Y., Chen, X., Liu, M., Gao, L., & Deng, H. (2020). A New Approach for Optimizing Rain Gauge Networks: A Case Study in the Jinjiang Basin. *Water*, *12*(8), Article 8. https://doi.org/10.3390/w12082252

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., & Mocko, D. (2012). 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). https://doi.org/10.1029/2011JD016048

Xu, H. (2006). Modification of Normalised Difference Water Index (ndwi) to Enhance Open Water Features in Remotely Sensed Imagery. *International Journal of Remote Sensing*, *27*(14), 3025–3033. https://doi.org/10.1080/01431160600589179

Yang, X., Zhao, S., Qin, X., Zhao, N., & Liang, L. (2017). Mapping of Urban Surface Water Bodies from Sentinel-2 MSI Imagery at 10 m Resolution via NDWI-Based Image Sharpening. *Remote Sensing*, *9*(6), Article 6. https://doi.org/10.3390/rs9060596

Yao, F., Wang, J., Wang, C., & Crétaux, J.-F. (2019). Constructing Long-Term High-Frequency Time Series of Global Lake and Reservoir Areas Using Landsat Imagery. *Remote Sensing of Environment*, *232*, 111210. https://doi.org/10.1016/j.rse.2019.111210

Zahid, M. N., Ahmad, S., Khan, J. A., Arshad, M. D., Azmat, M., & Ukasha, M. (2023). Evapotranspiration estimation using a satellite-based surface energy balance: A case study of Upper Bari Doab, Pakistan. *Environmental Earth Sciences*, *82*(24), 601. https://doi.org/10.1007/s12665-023-11284-5

Zawya News. (2024, March 14). *Crown Farms partners with University of Washington Civil & Environmental Engineering*. https://www.zawya.com/en/press-release/companies-news/crown-farms-partners-with-university-of-washington-civil-and-environmental-engineering-su3nvpf8

Zhang, C., Dong, J., & Ge, Q. (2022). IrriMap_CN: Annual irrigation maps across China in 2000–2019 based on satellite observations, environmental variables, and machine learning. *Remote Sensing of Environment*, *280*, 113184. https://doi.org/10.1016/j.rse.2022.113184

Zhang, H. K., Roy, D. P., Yan, L., Li, Z., Huang, H., Vermote, E., Skakun, S., & Roger, J.-C. (2018). Characterization of Sentinel-2A and Landsat-8 Top of Atmosphere, Surface, and Nadir BRDF Adjusted Reflectance and Ndvi Differences. *Remote Sensing of Environment*, *215*, 482–494. https://doi.org/10.1016/j.rse.2018.04.031

Zhang, H., Wang, G., Li, S., & Cabral, P. (2025). Understanding Evapotranspiration Driving Mechanisms in China With Explainable Machine Learning Algorithms. *Internation Journal of Climatology*. https://doi.org/10.1002/joc.8774

Zhu, Z., & Woodcock, C. (2012). Object-Based Cloud and Cloud Shadow Detection in Landsat Imagery. *Remote Sensing of Environment*, *118*, 83–94. https://doi.org/10.1016/j.rse.2011.10.028

Zoratipour, E., Mohammadi, A. S., & Zoratipour, A. (2023). Evaluation of SEBS and SEBAL algorithms for estimating wheat evapotranspiration (case study: Central areas of Khuzestan province). *Applied Water Science*, *13*(6), 137. https://doi.org/10.1007/s13201-023-01941-2

Appendix A

8.1. Creation of Configuration Files (Step 1):

The configuration files provide the essential information required for the sDRIPS system to tailor the advisory system to the specific needs of the user. There are three configuration files, namely - the Script Config File, Crop Config File, and Command Area Config File. These files store necessary file paths, information about the crops grown in each command area, the dates for which the system must run, and other pertinent details. Below are the details of the functionality of each of the configuration files.

Script Config File: This file stores the path to the command area shapefile, the date for which the system is to be run, and whether the user wants to run the system for the current week, last week, or both. Furthermore, script configuration file holds Boolean statements ("True" or "False") to consider different modules (say percolation or precipitation forecast) into the analysis

Crop Config File: This configuration file stores the crop coefficient (Kc) values as a function of days. By default, it includes Kc values for several crops. Having a configuration file for Kc values gives users the flexibility to update Kc values as needed, especially with the introduction of high-yielding crops or when a cropping pattern needs to be simulated for changing climate conditions. Additionally, the simple structure of the file allows users to add new crops and drop existing crops for water management analysis.

Command Area Config File: This file is created based on user inputs and is automatically scaled to each command area. It includes data such as the planting date, crop type, and soil coefficient (default is 0.5) for each command area. This setup provides users with the flexibility to input specific values for each command area. Additionally, users can set default values that will be automatically applied to all command areas, simplifying the configuration process.

In cases where users do not have the command area shapefile, sDRIPS can iteratively generate buffer regions around each canal until it reaches the required serving area. This approach requires information on how much area each canal serves. The method back calculates the command area boundary based on the canal's serving capacity. While this method may not be entirely accurate, it provides a reasonably precise boundary to begin with.

8.2. User Friendliness of sDRIPS

The sDRIPS system is designed to address the gaps identified in the current literature, as mentioned in section 1 of the study, by providing a user-friendly and customizable tool for analyzing various irrigation scenarios. The primary key features that enhance the usability of sDRIPS are as follows:

1. Ease of Use and Customization: sDRIPS employs configuration files that allow users to modify variables without altering the main script. To implement sDRIPS in a specific region, requirements are minimal, users need to have a Google Earth Engine (GEE) account and a Precipitation Processing System (PPS) account. Additionally, users must provide the canal network shapefiles, the command area of each canal (recommended but optional), the types of crops to be included in the analysis, and the planting dates.

2. Debugging Support: sDRIPS supports debugging by logging detailed information at every step. This includes completion reports, error messages, and the corresponding line numbers, which assist users in identifying and resolving issues efficiently.

3. Real-Time Analysis Status: During the execution of its algorithms, sDRIPS displays a progress bar for each algorithm, along with an estimated time to completion allowing users to monitor the status of ongoing processes.

4. Low Computational Power: sDRIPS leverages the GEE cloud computing platform for core processes such as Penman-Monteith and SEBAL-based ET estimation, thereby eliminating the need for high-performance computing resources. This ensures that users can run the system effectively without requiring advanced computational infrastructure.

8.3. Uncertainty Analysis

8.3.1. Through changing input parameters

We conducted Monte Carlo simulations by perturbing the six key input parameters, listed in Table 8.1. Two methods of perturbation were employed:

1. Applying fixed percentage changes based on satellite retrieval uncertainty, and

2. Randomly sampling values within realistic bounds informed by observational variability and literature.

Both methods yielded similar results in terms of ET distribution, reinforcing the robustness of the uncertainty propagation approach. Here we have focused on the approach (a).

Table 8.1: Parameters perturbed in Monte Carlo Simulations

Sr. No	Parameter	Range
1	Normalized Difference Vegetation Index (NDVI)	±20% from estimated
2	Albedo	±20% from estimated

3	Leaf Area Index (LAI)	±20% from estimated
4	Surface Roughness	±20% from estimated
5	Aerodynamic Resistance	±20% from estimated
6	Emissivity	±20% from estimated

Using these perturbed inputs, we generated an ensemble of 1,000 Monte Carlo simulations over the selected command area, covering multiple timestamps throughout the study period. The relative spread of each ensemble was calculated using **Error! Reference source not found.**, where Q_3 and Q_1 denote the 75th and 25th percentiles of the ET distribution, respectively. Figure 8.1 presents the boxplot of ET ensembles at a representative timestamp, demonstrating that the ensemble mean and median closely align with the deterministic (unperturbed) estimate for all time periods.

$$Realtive\ Spread = \frac{Q_3 - Q_1}{median}$$

(8.1)

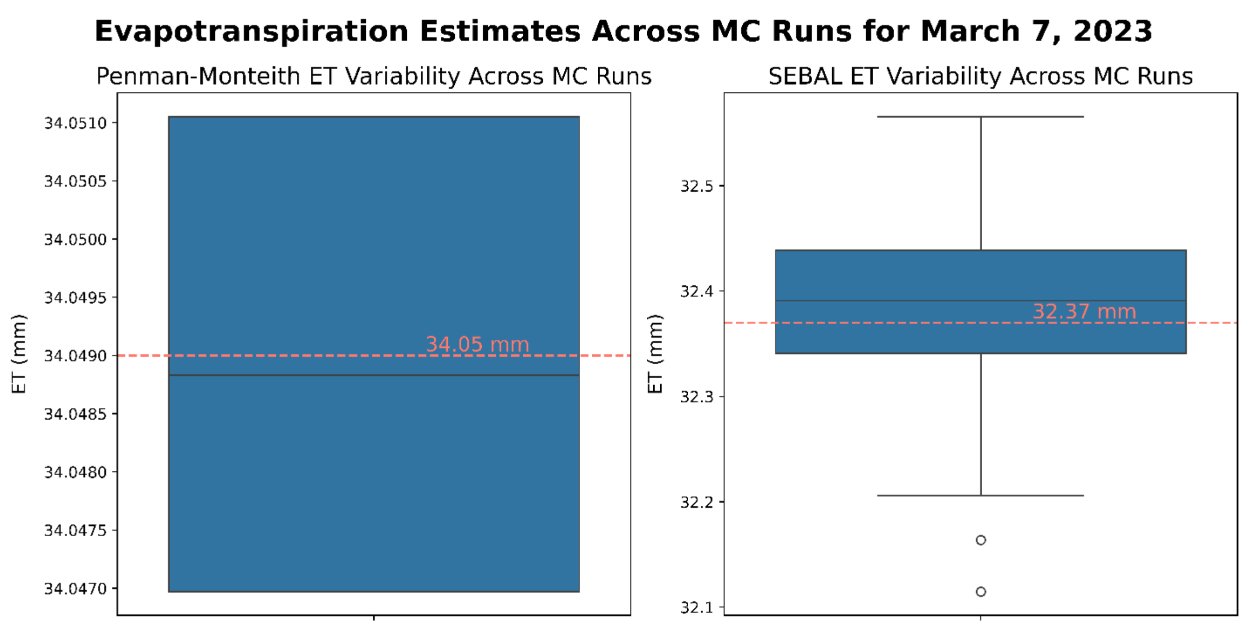

Figure 8.1: Box plots for ET estimates across 1000 Monte Carlo Simulations.

This non-parametric measure of uncertainty (relative spread) allowed us to generalize the uncertainty bounds to other TBP command areas with similar agro-climatic conditions. The mean relative spread was used as a scaling factor for extending uncertainty estimates region wide. Subsequently, the upper and lower bounds of the ET estimates were derived using **Error! Reference source not found.** and Equation (8.3)

, where ET_{det} is the deterministic or unperturbed estimate. The final uncertainty-propagated estimates for net water requirements are summarized in Table 8.2 and visualized in Figure 8.2.

$$Lower\ Bound = ET_{det} - \frac{Relative\ Spread\ \times ET_{det}}{2} \tag{8.2}$$

$$Upper\ Bound = ET_{det} + \frac{Relative\ Spread\ \times ET_{det}}{2} \tag{8.3}$$

Furthermore, the deterministic SEBAL-derived ET values were compared to the ensemble mean generated through the Monte Carlo simulations. Specifically, we computed the normalized deviation of the deterministic estimate from the ensemble mean using Equation (8.4),where SEBAL ET_{det} is the deterministic SEBAL estimate and $Ensemble_{std}$ is the standard deviation of the ensemble.

$$Normalized\ Deviation = \frac{SEBAL\ ET_{\mathrm{det}} - Ensemble_{mean}}{Ensemble_{std}} \times 100 \tag{8.4}$$

The mean normalized deviation over the operational period of the Teesta Barrage Project (January–April) was found to be 8.6%, indicating that the deterministic SEBAL values deviated by an average of 8.6% from the ensemble mean. This indicates that, on average, deterministic SEBAL estimates deviated by 8.6% from the ensemble mean, offering a quantitative measure of uncertainty. The relatively low deviation supports the validity of using deterministic SEBAL outputs as central estimates and justifies the scaling of results to other TBP command areas with similar climatic conditions.

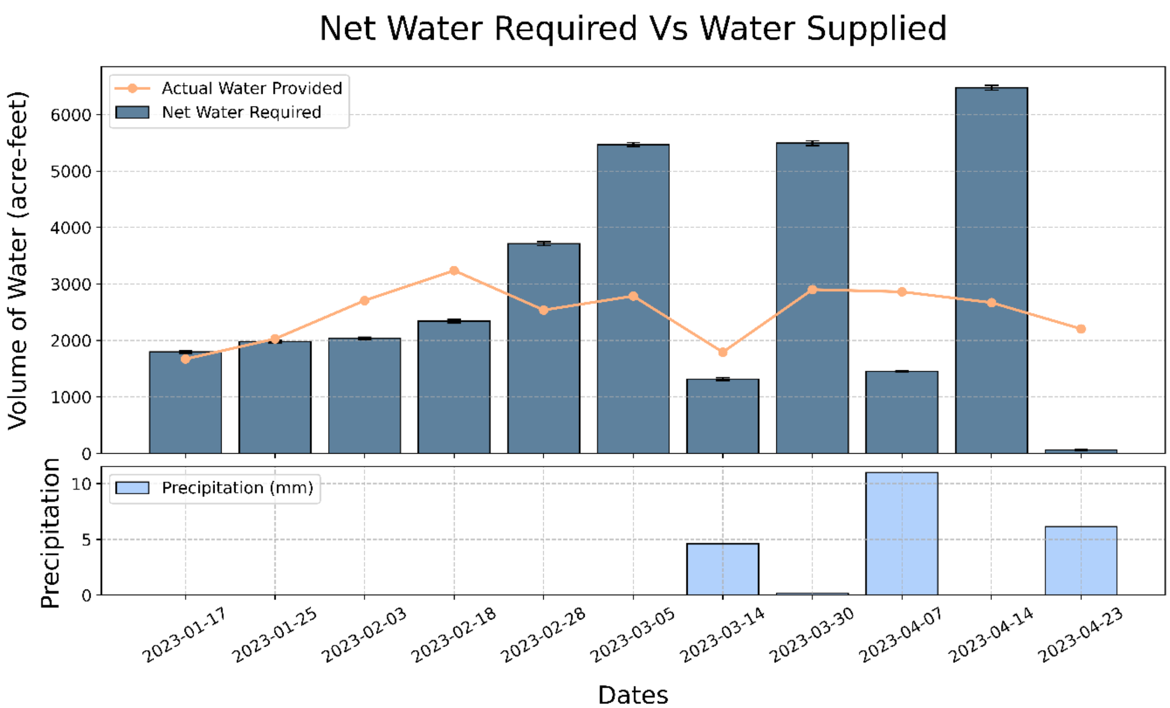

Figure 8.2: Timeseries of net water required by TBP along with the uncertainty bounds vs amount of water supplied to TBP.

Table 8.2: Net water required with uncertainty bounds.

Date	Net Water Required (acre-feet) (Lower Bound)	Net Water Required (acre-feet)	Net Water Required (acre-feet) (Upper Bound)	Bounds in Percentage (%)
2023-01-17	1776.25	1798.64	1821.02	1.24
2023-01-25	1960.36	1982.79	2005.21	1.13
2023-02-03	2017.60	2039.27	2060.94	1.06
2023-02-18	2316.17	2344.74	2373.32	1.22
2023-02-28	3682.03	3716.43	3750.83	0.93

2023-03-05	5436.24	5471.47	5506.70	0.64
2023-03-14	1290.84	1316.70	1342.70	1.96
2023-03-30	5455.14	5497.57	5540.00	0.77
2023-04-07	1441.59	1453.87	1466.14	0.84
2023-04-14	6436.03	6479.01	6521.99	0.66
2023-04-23	56.02	61.67	67.32	9.16

8.3.2. Through Sensitivity Analysis:

Sensitivity analysis focusing on two key drivers of evapotranspiration - air temperature and precipitation was conducted. According to Zhang et al., (2025) these parameters significantly influence evapotranspiration depending on the climatic regime. The analysis evaluates how uncertainties in these inputs impact evapotranspiration estimates and, consequently, the net water requirements.

1. Sensitivity to Air Temperature: We perturbed the air temperature by ±1°C and analyzed the response in both Penman-Monteith ET (potential ET under optimal water conditions) and SEBAL ET (actual ET under water stress). Figure 8.3 illustrates the results, where the upper panel presents the time series of seven-day cumulative Penman-Monteith evapotranspiration (mean of the region) and the lower panel displays the SEBAL evapotranspiration under the same perturbation. The expected trend is observed, with an increase in air temperature leading to a corresponding increase in evapotranspiration and a decrease in temperature resulting in lower evapotranspiration estimates. Furthermore, the temporal variation in evapotranspiration aligns with the crop growth cycle - initial, development, and harvesting phases. In terms of magnitude, a ±1°C change in air temperature leads to an approximate change of 1.4 mm in Penman-Monteith evapotranspiration and around 1 mm in SEBAL evapotranspiration, indicating that potential evapotranspiration under ideal conditions is slightly more sensitive to temperature variations compared to actual evapotranspiration under the water-stressed conditions.

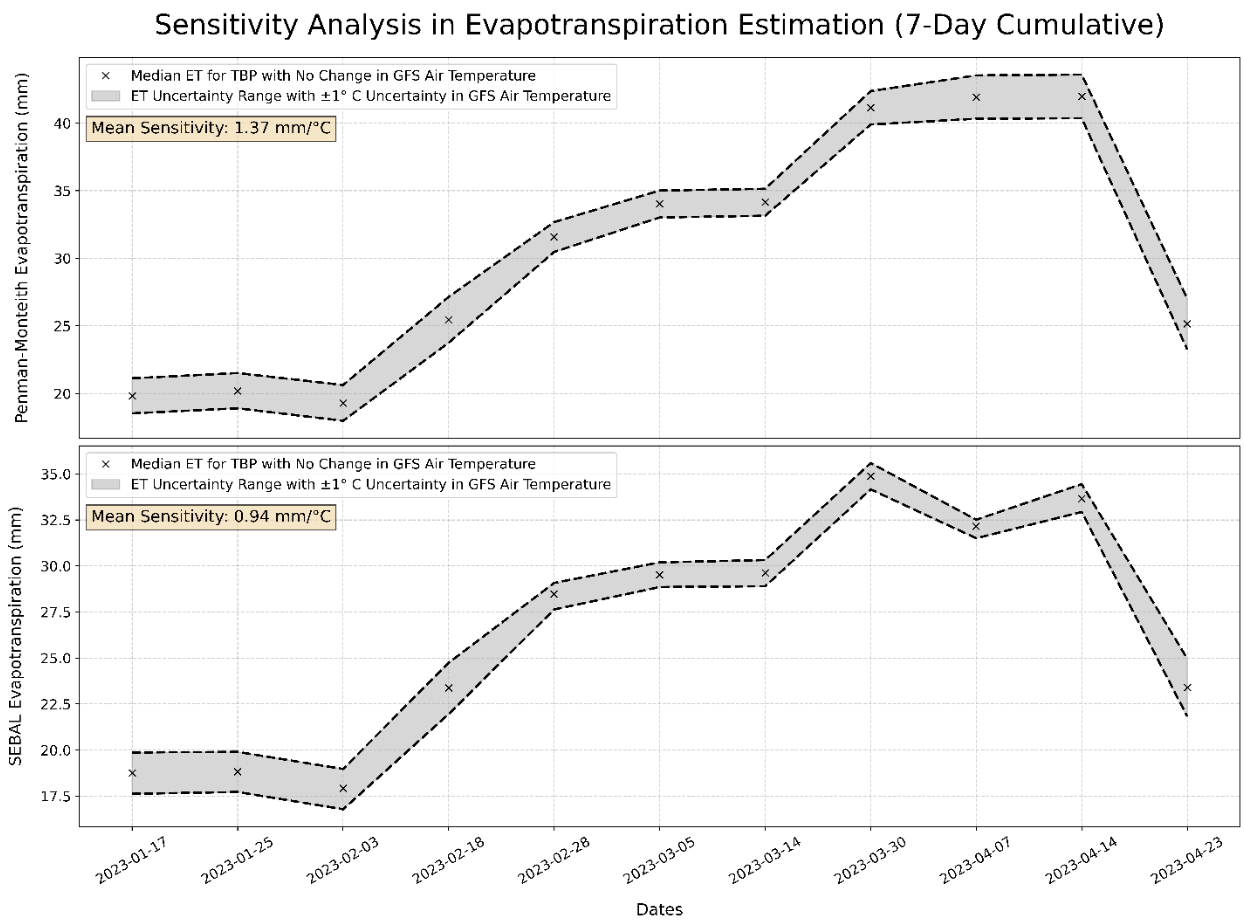

Figure 8.3: Sensitivity Analysis in Evapotranspiration Estimation

To further understand the implications of this variation, we assessed how these changes in evapotranspiration translate into changes in net water requirement. Figure 8.4 presents the sensitivity analysis in this context, consisting of two panels: the upper panel illustrates the volume of water required and the amount provided by in-situ supply, while the lower panel depicts the contribution of precipitation. The results indicate that a ±1°C change in air temperature leads to an approximately 0.4 million m³ variation in net water requirement, emphasizing the role of temperature fluctuations in water demand estimation.

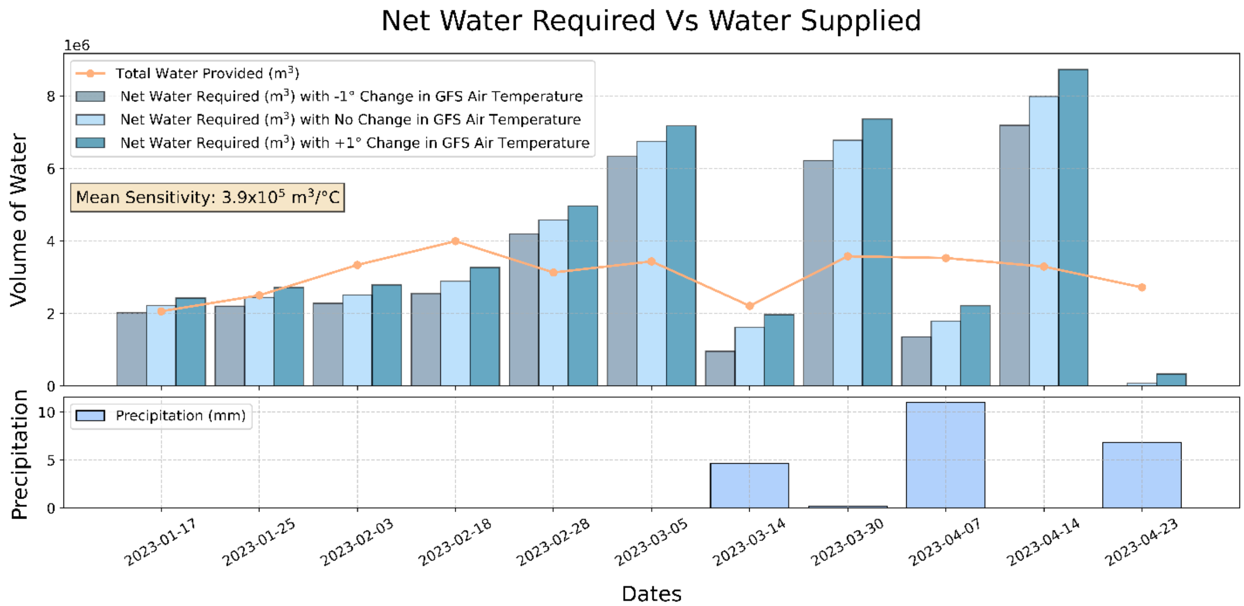

Figure 8.4: Sensitivity Analysis for Net Water Required w.r.t. Change in Air Temperature

2. Sensitivity to precipitation: Given its direct role in balancing water demand, we investigated the impact of a ±1 mm change in precipitation over the entire Teesta Barrage Project (TBP) region. Figure 8.5 presents this analysis, and the results indicate that a 1 mm increase in precipitation reduces the net water requirement by approximately 1 million m^3, demonstrating a more pronounced impact than air temperature fluctuations. This highlights the significance of precipitation uncertainty, as even minor deviations in precipitation estimates can substantially alter water demand calculations.

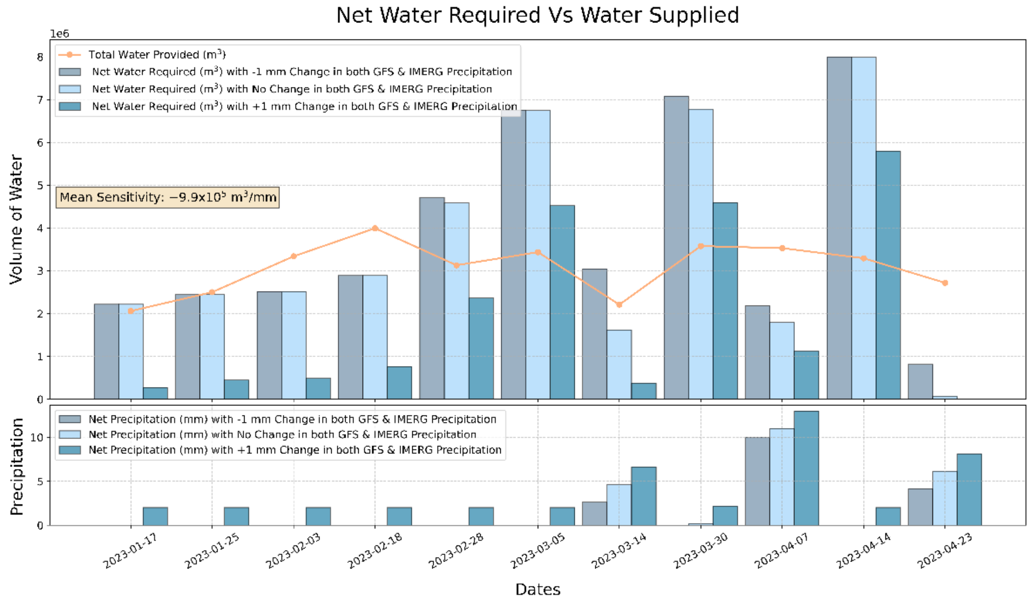

Figure 8.5: Sensitivity Analysis for Net Water Required w.r.t. Change in Precipitation

A comparison of these two factors suggests that in a hypothetical scenario, if water managers had control over modifying either air temperature or precipitation, increasing precipitation by 1 mm over entire TBP would be more beneficial than reducing air temperature by 1°C over entire TBP, as the former leads to greater water savings. This underscores the importance of accurate precipitation forecasting in improving water resource management and planning. Overall, our sensitivity analysis highlights the extent to which uncertainties in key meteorological inputs influence water demand estimations, emphasizing the need for robust data sources and improved forecasting methods to enhance the reliability of net water requirement assessments.

Appendix B

9.1. Configuration Setting for Use Case 1 – University of Washington

9.1.1. High-Level controls

High-Level controls from sdrips_config.yaml

```yaml
# -------------------------------
# HIGH-LEVEL MODULE CONTROLS
# -------------------------------
Command_Area_Net_Water_Requirement: true  # Generate advisory on how much water should be allotted to each command area
Canal_water_allotment: false          # Generate advisory on how much water should be allotted to each canal (From primary to tertiary canals)
Insitu_Sensor_integration: true  # Toggle to integrate in-situ sensor data with satellite data for bias correction
Weather_station_integration: false  # Toggle to integrate local weather station data to improve accuracy
```

9.1.2. Sensor-Level controls

Sensor-Level controls in sdrips_config.yaml

```yaml
# -------------------------------
# IN-SITU SENSOR CONFIGURATION
# -------------------------------
Insitu_Sensor_integration_config:
  air_temperature_sensor: true            # Use air temperature sensor data for correction. Eg - true or false
  air_temperature_variable: 'Temperature' # Reference name of the variable in insitu data. (If air_temperature_sensor is True)
  soil_moisture_sensor:  false            # Use soil moisture sensor data. Eg - true or false
  soil_moisture_variable: 'Water Content' # Reference name of the variable in insitu data. (If soil_moisture_sensor is True)
  wind_speed_sensor:  true               # Use wind speed sensor data. Eg - true or false
  wind_speed_variable: 'Wind'       # Reference name of the variable in insitu data. (If wind_speed_sensor is True)
  specific_humidity_sensor:  true           # Use specific humidity sensor data. Eg - true or false
  specific_humidity_variable: 'Specific Humidity'       # Reference name of the variable in insitu data. (If specific_humidity_sensor is True)
  hub_url: URL HAS BEEN REMOVED DUE TO PRIVACY  # URL to the sensor data hub (if applicable). Eg - "https://your_data_hub.com/Sensor_Cronjob/"
  date_regex: "(\\d{4}_\\d{2}_\\d{2})" # Regular expression to extract date from sensor filenames. 4 represents _ _ _ _ in Year, 2 represens _ _ in months and another 2 represents _ _ in date.
  date_format: "%Y_%m_%d" # Date format for sensor data filenames. Eg - "%Y_%m_%d" for 2024_04_06
  sensor_data_path:  Sensor_Data                    # Path to folder (should be in save_data_loc folder) containing sensor CSV files
  interpolation_method: best              # Options: 'best', 'average', 'idw', 'kriging', 'inverse_distance_weighting', 'nearest_neighbor'
```

9.1.3. Low-Level controls

Low-Level controls in sdrips_config.yaml

```
# -------------------------------------------------
# LOW-LEVEL MODULE CONTROLS (Common to every module)
# -------------------------------------------------
Irrigation_cmd_area_shapefile:
  path:  Shapefiles/UW_Boundary/UW_ROI.shp  # Path to the irrigation command area shapefile. Eg - Shapefiles/my_roi/my_roi.shp
  feature_name:   ADM3_EN    # Unique identifier for the irrigation command area features. Eg - ADM3_EN
  numeric_id_name:  id  # Numeric ID field for the irrigation command area features. Eg - ID or ADM3_EN (if a number)
  area_column_name: 'Irr_Area' # Column name in the canal shapefile that contains the area of the command area (in meters squared). Eg - 'Irrigabl_Area'

Irrigation_cmd_area_shapefile_Bounds:  # Bounds for the irrigation command area shapefile in EPSG:4326. If empty, sDRIPS consider bounds of the command area shapefile.
                                       # Here this param is given as the IMERG Precipitation spatial resolution is 25km and it might be possible
                                       # that the bounds need to be adjusted due to small size of command areas.
  leftlon:
  rightlon:
  toplat:
  bottomlat:

GEE_Asset_ID:   # GEE asset ID for the irrigation command area shapefile, if you have uploaded it to GEE, use 'id' otherwise use 'shp' for local shapefile.
                # If empty, sDRIPS will consider the local shapefile mentioned in 'Irrigation_cmd_area_shapefile'.
  id: # Path to the GEE asset. Eg - "users/skhan7/BWDB_Commad_Area"
  shp:  # Path to the local irrigation command area shapefile. Eg - Shapefiles/Irrigation_Command_Areas.shp

Date_Running:   # Date from which the sDRIPS will start running
  start_date:  "2025-08-18" # Eg - "2024-06-07"
  default_run_week: false # true - runs the sDRIPS for currentweek and lastweek. For manual run set it false and enter week in run_week below. By default - false
  run_week: ["currentweek"] # Enter if 'default_run_week' is set to false. Eg - ["currentweek", "lastweek"].

Save_Data_Location:    # Location where the output data will be saved
  save_data_loc: C:\Users\skhan7\OneDrive - UW\Desktop\Research\PhD\Chapter3\Testing_sDRIPS_Package\UW_Test\Data # Eg - Playground/Test/

Cmd_Area_Config:    # Configuration for the command area, giving info on how each command area should be processed
  path:  config_files/ca_config.yaml # Eg - config_files/uw_ca_config.yaml

Crop_Config:    # Configuration for the crop type, giving info on how each crop should be processed based on number of days after planting.
  path: config_files/crop_config.yaml

Config_links:
  path: config_files/config_links.yaml

Correction_iter:   # Number of iterations for the aerdoynamic correction process (By default, it is set to 3)
  correction_iter: 3

Multiprocessing:  # Number of cores to use for multiprocessing.
```

```
  cores: 7   # Set to null or leave out to use default cpu_count - 1

Clean_Directory:  # Whether to clear the output directory before running the sDRIPS
  clear_directory_condition: true

Run_ET_Estimation:  # Whether to run the ET estimation module
  et_estimation: true

Secrets_Path:       # Path to the secrets.yaml file, which contains sensitive information like IMERG credentials, GEE service account details.
  path: config_files/secrets.yaml

Download_ET_TIFF:   # Whether to download the ET TIFF files from GEE
  download_tiff: true

GLCC_Mask:        # Whether to use GLCC LULC for to the ET estimation or use your own LULC map
  glcc_mask: false
  glcc_id: "users/saswegee/shahzaib/Lidar_NDVI_Map_glcc"  # GEE asset ID for the GLCC LULC map

Percolation_Config:   # Whether to run the percolation estimation module to estimate net water requirement
  consider_percolation: true

Precipitation_Config: # Whether to consider precipitation to estimate net water requirement
  consider_preciptation: true

Weather_Config:  # Whether to download forecasted weather data for the ET estimation
  consider_forecasted_weather: true

Region_stats:   # Generates details CSV files with statistics for all the command area
  estimate_region_stats: true

Tiff_to_PNGs:  # Whether to convert the downloaded TIFF files to PNG format for easier visualization
  tiff_to_pngs: false
```

9.2. Configuration Setting for Use Case 2 – Teesta Barrage Project, Bangladesh

9.2.1. High-Level controls

```
# ------------------------------
# HIGH-LEVEL MODULE CONTROLS
# ------------------------------
Command_Area_Net_Water_Requirement: true  # Generate advisory on how much water should be allotted to each command area
Canal_water_allotment: false          # Generate advisory on how much water should be allotted to each canal (From primary to tertiary canals)
Insitu_Sensor_integration: false  # Toggle to integrate in-situ sensor data with satellite data for bias correction
Weather_station_integration: false  # Toggle to integrate local weather station data to improve accuracy
```

9.2.2. Low-Level controls

Low-Level controls in sdrips_config.yaml file

```
# --------------------------------------------------
# LOW-LEVEL MODULE CONTROLS (Common to every module)
# --------------------------------------------------
```

```yaml
Irrigation_cmd_area_shapefile:
  path: Shapefiles\TBP_Command_Area\commandarea_teesta_ph1_simplified_modified_Irrigable_Area_Final.shp  # Path to the irrigation command area shapefile. Eg - Shapefiles/my_roi/my_roi.shp
  feature_name: CNLNM_ID    # Unique identifier for the irrigation command area features. Eg - ADM3_EN
  numeric_id_name: ID   # Numeric ID field for the irrigation command area features. Eg - ID or ADM3_EN (if a number)
  area_column_name: Irrigabl_1 # Column name in the canal shapefile that contains the area of the command area (in meters squared). Eg - 'Irrigabl_Area'

Irrigation_cmd_area_shapefile_Bounds:  # Bounds for the irrigation command area shapefile in EPSG:4326. If empty, sDRIPS consider bounds of the command area shapefile.
                                       # Here this param is given as the IMERG Precipitation spatial resolution is 25km and it might be possible
                                       # that the bounds need to be adjusted due to small size of command areas.
  leftlon:
  rightlon:
  toplat:
  bottomlat:

GEE_Asset_ID:   # GEE asset ID for the irrigation command area shapefile, if you have uploaded it to GEE, use 'id' otherwise use 'shp' for local shapefile.
                # If empty, sDRIPS will consider the local shapefile mentioned in 'Irrigation_cmd_area_shapefile'.
  id: # Path to the GEE asset. Eg - "users/skhan7/BWDB_Commad_Area"
  shp: Shapefiles\TBP_Command_Area\commandarea_teesta_ph1_simplified_modified_Irrigable_Area_Final.shp # Path to the local irrigation command area shapefile. Eg - Shapefiles/Irrigation_Command_Areas.shp

Date_Running:   # Date from which the sDRIPS will start running
  start_date: "2023-02-28"  # Eg - "2024-06-07"
  default_run_week: false # true - runs the sDRIPS for currentweek and lastweek. For manual run set it false and enter week in run_week below. By default - false
  run_week: ["currentweek"] # Enter if 'default_run_week' is set to false. Eg - ["currentweek", "lastweek"].

Save_Data_Location:    # Location where the output data will be saved
  save_data_loc: C:\Users\skhan7\OneDrive - UW\Desktop\Research\PhD\Chapter3\Testing_sDRIPS_Package\Full_Run_TBP\Data # Eg - Playground/Test/

Cmd_Area_Config:    # Configuration for the command area, giving info on how each command area should be processed
  path: config_files/ca_config.yaml # Eg - config_files/uw_ca_config.yaml

Crop_Config:    # Configuration for the crop type, giving info on how each crop should be processed based on number of days after planting.
  path: config_files/crop_config.yaml

Config_links:
  path: config_files/config_links.yaml

Correction_iter:   # Number of iterations for the aerdoynamic correction process (By default, it is set to 3)
  correction_iter: 3

Multiprocessing:  # Number of cores to use for multiprocessing.
  cores:    # Set to null or leave out to use default cpu_count - 1

Clean_Directory:  # Whether to clear the output directory before running the sDRIPS
  clear_directory_condition: true

Run_ET_Estimation:  # Whether to run the ET estimation module
```

```yaml
  et_estimation: true

Secrets_Path:       # Path to the secrets.yaml file, which contains sensitive information like IMERG credentials, GEE service account details.
  path: config_files/secrets.yaml

Download_ET_TIFF:   # Whether to download the ET TIFF files from GEE
  download_tiff: true

GLCC_Mask:         # Whether to use GLCC LULC for to the ET estimation or use your own LULC map
  glcc_mask: true
  glcc_id:    # GEE asset ID for the GLCC LULC map

Percolation_Config:   # Whether to run the percolation estimation module to estimate net water requirement
  consider_percolation: true

Precipitation_Config: # Whether to consider precipitation to estimate net water requirement
  consider_preciptation: true

Weather_Config:  # Whether to download forecasted weather data for the ET estimation
  consider_forecasted_weather: true

Region_stats:   # Generates detailed CSV files with statistics for all the command area
  estimate_region_stats: true

Tiff_to_PNGs:  # Whether to convert the downloaded TIFF files to PNG format for easier visualization
  tiff_to_pngs: false
```