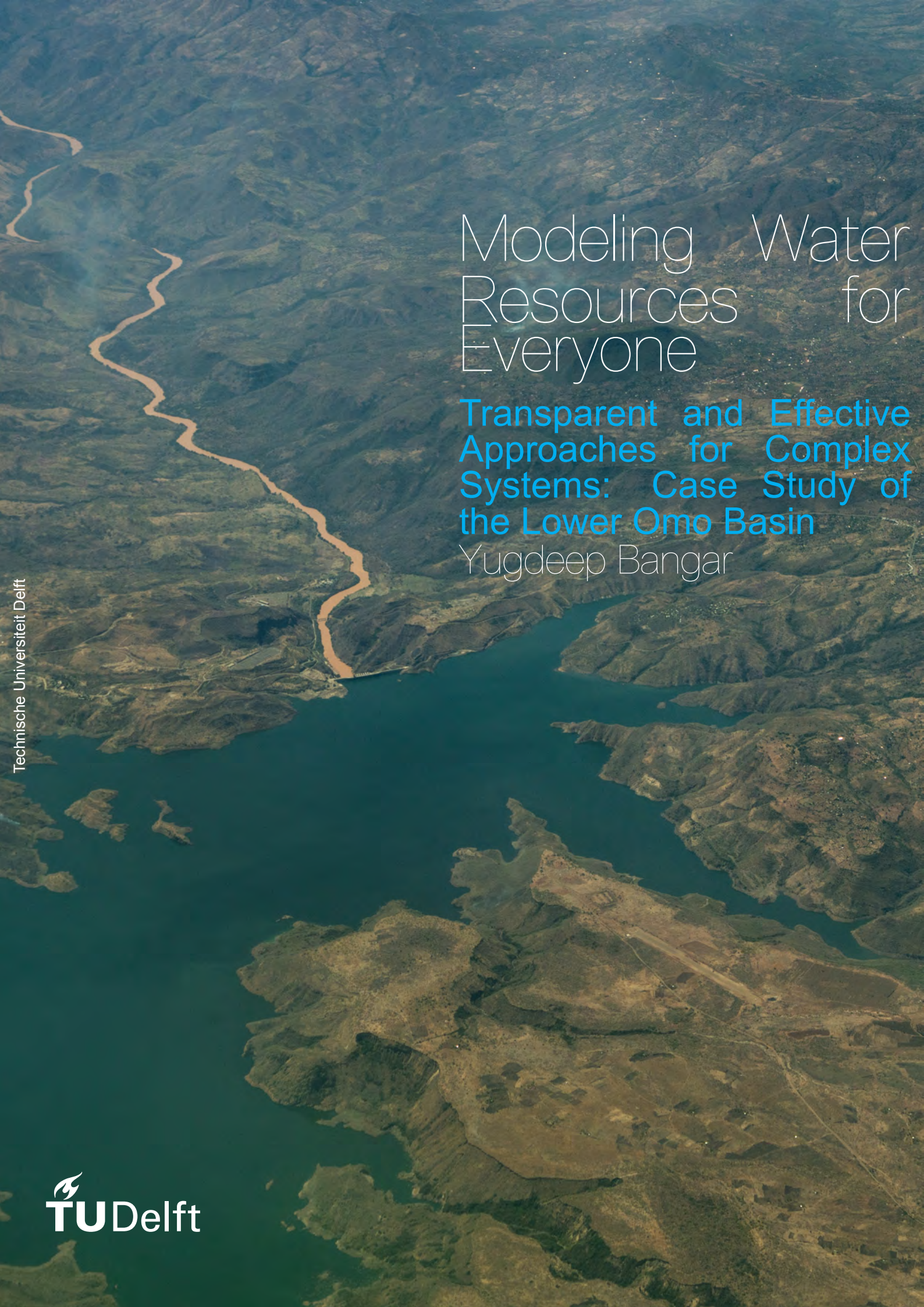


Modeling Water Resources for Everyone

Transparent and Effective Approaches for Complex Systems: Case Study of the Lower Omo Basin

Yugdeep Bangar

Technische Universiteit Delft

TU Delft

Modeling Water Resources for Everyone

Transparent and Effective Approaches for Complex Systems: Case Study of the Lower Omo Basin

by

Yugdeep Bangar

to obtain the degree of Master of Science
in Engineering and Policy Analysis
at the Delft University of Technology.

Student number:	4733967	
Thesis committee:	Dr. P. Bots,	TU Delft, Chairperson
	Dr. J. Salazar,	TU Delft, First Supervisor
	Prof. P. van Gelder,	TU Delft, Second Supervisor
	Dr. S. Yalew,	IHE Delft, External Supervisor

The Python implementation of HydroWizard Framework is available on following links:
Package: `https://pypi.org/project/hydrowizard/`
Documentation: `https://hydrowizard.readthedocs.io/`
Source code: `https://github.com/yugdeep/hydrowizard`

Data and code for the Case Study on Lower Omo Basin using HydroWizard package is available at `https://github.com/yugdeep/omo-gibe`.

An electronic version of this thesis is available at `https://repository.tudelft.nl/`.

TU Delft

Acknowledgments

This thesis represents not just the culmination of my research but also the collaborative effort and support of many individuals to whom I owe my deepest gratitude.

I am profoundly grateful to my supervisor, Dr. Jazmin Zatarain Salazar, whose expertise and mentorship has been pivotal to this research. Her insights into Evolutionary Multi-Objective Direct Policy Search (EMODPS) provided valuable direction for this work, while her astute advice on project scope ensured a balance between ambition and feasibility. Dr. Salazar's efforts to connect me with the broader research community have been crucial, broadening the perspectives incorporated into this study. Her enthusiasm and dedication to scientific rigor have been truly motivating throughout the process.

Equally, I extend my heartfelt thanks to Dr. Pieter Bots, my thesis chairperson, whose intellectual rigor and unwavering support have been instrumental throughout this process. Our thought-provoking discussions challenged me to think critically and explore new perspectives, significantly enriching this work. Dr. Bots' probing questions and constructive feedback have improved this thesis and broadened my approach to problem-solving in water resource management. His enduring confidence in my abilities has been a driving force, pushing me to exceed my expectations.

I sincerely appreciate Prof. Pieter van Gelder's exceptional motivational support and encouragement. His belief in this research's potential has been a driving force, enabling me to refine and seamlessly integrate the technical and development work with the underlying research objectives. His enthusiasm for its collaborative potential and guidance was instrumental in shaping the project into a form that met academic standards and held promise for real-world impact.

A special note of thanks goes to Dr. Seleshi Yalew for providing crucial region-specific and case-specific data for the Lower Omo Basin. His expertise and willingness to share information have been vital in grounding this research in real-world contexts and challenges.

This thesis was only possible with the collaborative environment fostered by the faculty and my fellow TU Delft and IHE Delft researchers. The academic rigor and innovative spirit at TU Delft have profoundly shaped my research approach.

Lastly, I want to thank my family and friends for their patience, understanding, and unwavering support throughout this academic journey. Their encouragement has been my bedrock, providing the stability needed to complete this work.

To all who have contributed to this journey, directly or indirectly, I offer my heartfelt gratitude. This thesis is as much a product of your support as it is of my efforts.

Cover Image Credit: Aerial Landscape view of area around the "Gigel Gibe III Dam" on Omo River located in southwest Ethiopia by Mario Hagen. Source: Adobe Stock.

Executive Summary

The global challenge of sustainable water resource management has never been more pressing, with rapidly developing regions facing the complex task of balancing competing demands amidst data scarcity and environmental uncertainties. This thesis targets a critical void in our approach to these complex issues: the lack of efficient, transparent, reproducible, and extensible modeling frameworks capable of navigating these multifaceted challenges. At its core, the research is driven by the question: *"How can we transparently and efficiently model a water resource system and search for optimal policies for competing water uses, considering uncertainties from evolving development projects and data scarcity, as exemplified by the Lower Omo Basin?"*

To answer this, the study introduces HydroWizard, an innovative framework designed to revolutionize water resource modeling. HydroWizard represents a paradigm shift in approaching complex water resource challenges, seamlessly blending scientific rigor with practical applicability. It empowers researchers, policymakers, and water resource managers to navigate the intricate landscape of 21st-century water management with unprecedented clarity and efficiency.

The framework's architecture is built on two pivotal components: a flexible, YAML-based Model Specification Language and a sophisticated Model Execution Engine. The specification language serves as a transparent, human-readable interface for defining complex water systems, encompassing nodes, flows, constraints, and objectives. This approach enhances model clarity and facilitates collaboration among diverse stakeholders, from technical experts to policymakers. The execution engine, leveraging advanced algorithms and a Directed Acyclic Graph (DAG) representation, translates these specifications into computationally efficient simulations and optimizations. This duality ensures that HydroWizard can handle the complexity of real-world water systems while maintaining accessibility and transparency.

The concrete realization of HydroWizard as a full-fledged Python package brings the framework's innovative concepts to life, making them accessible to a wide range of users. This implementation provides command-line and Python API interfaces out of the box, while its flexible design allows for easy expansion to graphical interfaces. The package's modular architecture and built-in support for parallel computing enable seamless handling of large-scale systems and easy integration of new components, pushing the boundaries of what is possible in water resource modeling. To ensure widespread adoption and collaborative development, the HydroWizard package is publicly available through PyPI, supported by extensive documentation and open-source code. This open approach not only enhances the tool's accessibility but also fosters a growing community of users and developers, driving continuous improvement and innovation in water resource management practices. HydroWizard paves the way for more efficient, transparent, and adaptive approaches to tackling complex water resource challenges by providing both a robust conceptual framework and its practical implementation.

To demonstrate the framework's capabilities and address the overarching research question, this study applies HydroWizard to the Lower Omo-Gibe River Basin in Ethiopia. This region epitomizes the challenges faced in many developing areas, with its intricate balance of hydropower generation, extensive irrigation demands, and critical environmental flow requirements. The model incorporates the Gibe III and Koysha dams, irrigation schemes for the Kuraz sugar plantations, and environmental flow requirements, capturing the basin's complex dynamics under data scarcity and rapid development conditions.

The research employs a sophisticated Evolutionary Multi-Objective Direct Policy Search (EMODPS) methodology, integrated with Radial Basis Function (RBF) networks, to navigate the multidimensional decision space of reservoir operations. Through an exhaustive optimization process comprising ten independent runs—each initialized with a unique random seed and executing 80,000 function evaluations—the study identifies a comprehensive set of 283 Pareto-optimal policies. These policies collectively define the efficient frontier of water resource management strategies, with each policy embodying a distinct equilibrium among three critical objectives: maximization of hydropower generation, optimization of irrigation demand fulfillment, and preservation of environmental flows. This rigorous multi-objective optimization framework offers decision-makers an intricate understanding of the inherent trade-offs in

water resource allocation, thereby facilitating evidence-based and balanced policy formulation in the context of competing water demands.

The results reveal nuanced insights into water management trade-offs. Policies optimized for irrigation eliminate monthly demand deficits but at the cost of reducing environmental flows by up to 48%. Conversely, environmentally-focused policies maintain optimal river flows but compromise irrigation capacity, potentially increasing irrigation deficits to 51%. Intriguingly, the study uncovers a relative consistency in mean power generation across these varied policies, suggesting the possibility of maintaining energy outputs while addressing other critical objectives. These findings highlight the complex interplay between different water uses and the importance of considering multiple objectives in water resource planning.

HydroWizard introduces groundbreaking visualization techniques that transform complex data into intuitive insights. Animated rule curves, derived from RBF networks, offer a dynamic representation of policy decisions across varying system states. These visualizations allow stakeholders to understand how reservoir operations change in response to different conditions, such as varying water levels or seasonal fluctuations. Complemented by detailed system state graphs, these tools provide unprecedented clarity in understanding and validating model behavior, bridging the gap between sophisticated algorithms and practical decision-making.

The framework's versatility is further demonstrated through its successful application to diverse water systems beyond the Lower Omo-Gibe, including the expansive Zambezi River Basin and a hypothetical minimal basin. This broader application underscores HydroWizard's potential as a universal tool for water resource modeling, adaptable to a wide range of contexts and scales, from small watershed management to transboundary river basin planning.

Key contributions of this research include:

- Development of the HydroWizard framework, offering a new standard in transparent, efficient, and reproducible water resource modeling.
- Implementation and public release of HydroWizard as an open-source Python package, democratizing access to advanced water resource modeling tools.
- Application of the framework to the Lower Omo-Gibe Basin, providing valuable insights for sustainable water management in a critical region.
- Introduction of novel visualization techniques, including animated rule curves and system state graphs, enhancing the interpretability of complex water management strategies.
- Demonstration of the framework's versatility through application to multiple river basins, showcasing its potential for widespread adoption in diverse contexts.

While marking significant progress, this study also identifies avenues for future enhancement. Potential improvements include incorporating more comprehensive uncertainty and sensitivity analyses, especially considering climate change and socio-economic dynamics. Future work could integrate advanced climate models and uncertainty quantification methods, bridging the gap between technical analysis and practical application. These enhancements would build upon the current framework, offering decision-makers even more robust tools for addressing complex water management challenges.

In conclusion, HydroWizard represents a significant leap forward in water resource modeling, offering a solution to the long-standing challenges of efficiency, transparency, reproducibility, and extensibility. By providing an open-source, accessible tool for complex water system analysis, this research contributes valuable insights for the Lower Omo-Gibe Basin. It sets a new standard for global water resource management studies. As we face increasing challenges in water resource management, exacerbated by climate change and growing demands, HydroWizard emerges as an innovative approach to these pressing issues.

This thesis marks an important step in our journey towards sustainable, equitable, and resilient water futures. It offers a path forward where complex scientific modeling becomes more accessible, where clear, visual insights inform decision-making, and where the global community of water resource professionals can collaborate more effectively. As water scarcity and management challenges intensify worldwide, the tools and approaches developed in this research provide a robust foundation for more informed, transparent, and sustainable water resource management, contributing to a future where water needs are balanced with environmental sustainability and social equity.

Contents

List of Figures

List of Tables

1

Introduction

Water, the lifeblood of our planet, stands at the nexus of human development, environmental sustainability, and global security. As we progress further into the 21st century, the management of this vital resource has become increasingly complex, demanding innovative approaches that can navigate the intricate web of competing demands, environmental uncertainties, and rapidly evolving socio-economic landscapes (Vörösmarty et al., 2000; IPCC, 2014). This thesis, titled “Modeling Water Resources for Everyone” addresses a critical gap in the field of water resource management by introducing a novel framework that promises to revolutionize how we model, analyze, and make decisions about water resources.

The global challenge of water resource management is multifaceted. On the one hand, burgeoning populations and expanding economies are placing unprecedented demands on water resources (Oki and Kanae, 2006). Conversely, climate change is altering precipitation patterns, increasing the frequency of extreme weather events, and introducing new levels of uncertainty into water availability projections (Milly et al., 2008). These challenges are particularly acute in rapidly developing regions, where the imperative for economic growth often clashes with environmental conservation and traditional water use practices (Biswas, 2004).

Amidst these challenges, the need for robust, transparent, and accessible water resource modeling tools has never been more pressing. Traditional approaches to water resource modeling often suffer from several limitations. Many existing models operate as “black boxes,” making it difficult for stakeholders to understand and trust the decision-making processes (Jakeman et al., 2006). The use of proprietary software or complex, poorly documented code hampers the scientific principle of reproducibility (Hutton et al., 2016). Furthermore, many models are tailored to specific contexts and need extensive work to adapt to new scenarios or incorporate emerging data sources (Pahl-Wostl, 2007). Advanced modeling techniques often require specialized expertise, limiting their use by a broader range of stakeholders (Voinov et al., 2016).

This thesis addresses these limitations head-on by introducing HydroWizard, an innovative framework designed to democratize water resource modeling. HydroWizard represents a paradigm shift in approaching complex water resource challenges, seamlessly blending scientific rigor with practical applicability. At its core, HydroWizard is built on two pivotal components: a flexible, YAML-based Model Specification Language that serves as a transparent, human-readable interface for defining complex water systems and a sophisticated Model Execution Engine that leverages advanced algorithms and a Directed Acyclic Graph (DAG) representation to translate these specifications into computationally efficient simulations and optimizations. This duality ensures that HydroWizard can handle the complexity of real-world water systems while maintaining accessibility and transparency for a wide range of users, from technical experts to policymakers.

To demonstrate the power and versatility of this new approach, we apply HydroWizard to a case study of the Lower Omo-Gibe River Basin in Ethiopia. This region exemplifies the complex challenges faced in many developing areas, with its intricate balance of hydropower generation, extensive irrigation demands, and critical environmental flow requirements (Hodbod et al., 2019; Avery and Tebbs, 2018). The Lower Omo-Gibe Basin is home to ambitious infrastructural developments, including the recently constructed Gibe III dam and the ongoing Koysha dam project (Zaniolo et al., 2021c). While these

projects promise significant economic benefits through energy production and agricultural expansion, they also pose substantial risks to downstream ecosystems and traditional livelihoods (Avery, 2017).

The Lower Omo-Gibe Basin case study serves as an ideal testing ground for HydroWizard. It demonstrates how the framework can efficiently model complex water systems with multiple reservoirs, diverse water uses, and competing stakeholder objectives. We show how HydroWizard handles data scarcity and uncertainty scenarios, which are common in many developing regions. The framework provides transparent and interpretable results that can inform policy decisions and stakeholder negotiations. Moreover, it facilitates the exploration of trade-offs between different water management objectives through advanced multi-objective optimization techniques.

1.1. Research Problem

Despite extensive research on the hydrological and socio-economic impacts of water resource developments, there remains a significant gap in the transparent and efficient modeling of complex water resource systems. This gap is particularly evident when considering competing water uses and the uncertainties arising from evolving development projects and data scarcity. Current water resource models often involve custom simulation and optimization code with hardcoded relationships, which significantly affects the models' reproducibility, transparency, and extensibility. Moreover, there is a notable lack of a general-purpose modeling framework that integrates advanced optimization techniques like the Evolutionary Multi-Objective Direct Policy Search (EMODPS) for water resource systems.

1.2. Research Objectives

This research aims to address these gaps by developing and applying a transparent, reproducible framework for water resource modeling that integrates EMODPS to the Lower Omo-Gibe River Basin. Specifically, this study seeks to:

1. Conduct a comprehensive stakeholder analysis to identify and quantify the main water usage objectives in the Lower Omo-Gibe River Basin.
2. Develop and implement the HydroWizard framework, featuring a transparent model specification language and an efficient execution engine capable of handling complex water systems.
3. Apply the HydroWizard framework to create a comprehensive model of the Lower Omo-Gibe Basin, incorporating available data and developing strategies to address data gaps.
4. Employ advanced multi-objective optimization techniques within the HydroWizard framework to identify and analyze Pareto-optimal policy choices for water resource management in the basin.

1.3. Research Questions

The overarching research question guiding this study is:

> *How can we transparently and efficiently model a water resource system and search for optimal policies for competing water uses, considering uncertainties from evolving development projects and data scarcity, as exemplified by the Lower Omo Basin?*

To address this central question, we break it down into four sub-questions:

1. What are the principal objectives concerning water usage among key stakeholders in the Lower Omo-Gibe River Basin?
2. How can we model a water resource system in a transparent, modular, and efficient manner, considering the complexities of the Lower Omo-Gibe River Basin?
3. What model configuration can be used for the Lower Omo-Gibe River Basin to analyze the trade-offs among water uses, given the data scarcity challenges?
4. What compromises are revealed by Pareto-optimal policy choices for water resource management in the Lower Omo-Gibe River Basin?

By addressing these questions and objectives, this research aims to make several significant contributions to the field of water resource management. We introduce a novel, open-source framework (HydroWizard) that sets a new standard for transparent, efficient, and reproducible water resource modeling. We develop innovative visualization techniques, including animated rule curves and system state graphs, to enhance the interpretability of complex water management strategies. Through the application of advanced multi-objective optimization techniques, we reveal nuanced insights into water management trade-offs in the Lower Omo-Gibe Basin. Finally, we demonstrate how complex modeling and optimization techniques can be made accessible to a broader range of stakeholders, potentially transforming decision-making processes in water resource management.

1.4. Thesis Structure

This thesis is organized into seven chapters, each addressing a crucial aspect of the research:

Chapter 1, this Introduction, presents the research context, problem statement, research questions, and objectives and provides an overview of the thesis structure.

Chapter 2, Background and Literature Review, offers a comprehensive overview of the Lower Omo-Gibe River Basin and a critical review of existing literature on water resource management practices, modeling techniques, and stakeholder objectives. This chapter also identifies critical research gaps that this study aims to address.

Chapter 3, Hydrological Modelling of the Lower Omo-Gibe Basin, details the data collection and processing methods used to model the basin, including strategies for handling data scarcity. It also discusses the assumptions and limitations of the hydrological modeling approach.

Chapter 4, Methodology, introduces the theoretical foundations of the HydroWizard framework after giving a background on the XLRM conceptual framework and the Evolutionary Multi-Objective Direct Policy Search (EMODPS) methodology. It provides a detailed description of HydroWizard's architecture, including its model specification language, execution engine, optimization module, and visualization tools.

Chapter 5, Results, presents the findings of the study. This includes a convergence analysis of the optimization algorithm, a Pareto front analysis revealing trade-offs between different water management objectives, policy analysis using novel visualization techniques (including animated rule curves derived from RBF networks), and an examination of the physical implications of different policies on reservoir storage, power generation, and water allocation.

Chapter 6, Discussion, synthesizes the study's findings, addressing the research questions and exploring broader implications. It examines the HydroWizard framework's methodological innovations and technical advancements, analyzes the trade-offs revealed in the Lower Omo-Gibe Basin study, and discusses policy implications and contributions to water resource management.

Chapter 7, Conclusion, summarizes key contributions, reflects on HydroWizard's impact, discusses implications for water security, outlines future research, and emphasizes the study's significance for sustainable water management.

Through this structured approach, the thesis aims to comprehensively explore how innovative modeling techniques can transform water resource management, making it more transparent, efficient, and accessible to a broader range of stakeholders. By bridging the gap between complex scientific modeling and practical decision-making, this research contributes to the broader goal of achieving sustainable and equitable water resource management in an increasingly water-stressed world.

2

Background and Literature Review

This chapter provides a comprehensive overview of the Lower Omo-Gibe River Basin and the current state of water resource management research. It begins with a detailed study area description, including its geographical, hydrological, and socio-economic characteristics. The chapter then examines the perspectives of key actors in the basin, namely Ethiopian Electric Power, the Ethiopian Sugar Industry Group, and local communities. An extensive literature review is presented, encircling hydrological modeling, ecological impacts, socio-economic consequences, and the application of advanced optimization techniques in water resource management. The chapter concludes by identifying critical research gaps that this study aims to address.

2.1. Study Area

The Lower Omo Basin, situated in southern Ethiopia, is a region of profound hydrological and ecological significance. Spanning approximately 79,000 km², the basin extends from the confluence of the Omo and Gibe rivers to the Turkana Basin in Kenya (Avery and Tebbs, 2018). This area is characterized by a diverse range of ecosystems, including highland forests, savannas, and wetlands, contributing to its rich biodiversity (Hodbod et al., 2019).

Geographically, the basin is located between 4°30'N to 9°00'N latitude and 35°00'E to 38°00'E longitude. It is bordered by the Ethiopian highlands to the north and Lake Turkana to the south (Zaniolo et al., 2021c). The Omo River, the principal watercourse of the basin, originates in the Ethiopian highlands and flows southwards for about 760 kilometers before emptying into Lake Turkana (Carr, 2017).

The climate of the Lower Omo Basin exhibits significant variation with altitude. Highland areas experience a temperate climate, while lowland regions are characterized by arid to semi-arid conditions (Chaemiso et al., 2016). Annual rainfall in the basin ranges from 400 mm in the lowlands to 1,900 mm in the highlands, with mean annual temperatures varying from less than 17°C in the highlands to over 29°C in the lowlands (Dagne et al., 2023). This climatic diversity plays a crucial role in shaping the hydrological regime of the basin and supporting its varied ecosystems.

The socio-economic fabric of the Lower Omo Basin is equally diverse, comprising various ethnic groups engaged in pastoralism, agro-pastoralism, and flood-retreat cultivation (Amos et al., 2021). The basin is crucial for the livelihoods of these communities, who rely heavily on the seasonal flooding of the Omo River for their agricultural activities (Hodbod et al., 2019). Traditional practices such as flood-retreat agriculture and fishing have been sustained for generations, forming an integral part of the region's cultural heritage (Carr, 2017).

Figure 2.1 offers an overview of the Omo-Turkana Basin's geographical context. The map captures the basin's role as a shared resource across multiple East African nations. It eloquently illustrates the far-reaching implications of water resource decisions in this area by situating the basin within its broader African context.

Figure 2.2 presents a detailed digital elevation map of the Omo-Turkana Basin in Ethiopia. It depicts the basin's topographic variability and the extent of water resource infrastructure development. The map provides a spatial context for understanding the distribution of water resources and the potential impacts of infrastructure projects on different parts of the basin.

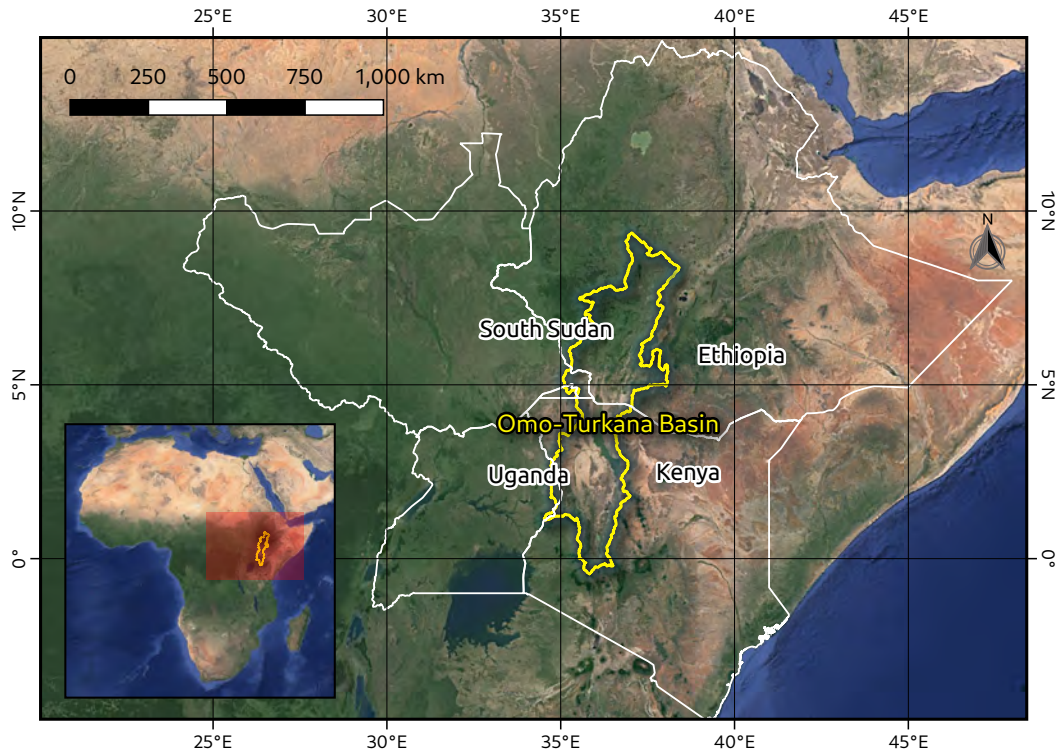


Figure 2.1: Geographical location of the Omo-Turkana Basin in East Africa. This map illustrates the transboundary nature of the basin, spanning parts of Ethiopia, Kenya, Uganda, and South Sudan. The inset map provides context for the basin's location within the African continent, emphasizing its significance in the Horn of Africa region.

Figure 2.3 provides a schematic representation of the main components of the Omo River Basin. It distills the complex hydrological system into a clear, comprehensible format. The schematic diagram effectively captures our research focus by highlighting the study area within the larger system, providing a valuable reference point.

2.2. Actor Perspective

The Lower Omo-Gibe River Basin is characterized by a complex interplay of diverse stakeholders, each with their own objectives, interests, and impacts on water resource management. This section examines the perspectives of three key actors in the basin: Ethiopian Electric Power (EEP), the Ethiopian Sugar Industry Group (ESIG), and the local communities. Understanding these different viewpoints is crucial for developing a comprehensive and balanced approach to water resource management in the region.

By analyzing each actor's goals, activities, and concerns, we can better appreciate the competing demands on the basin's water resources and the potential conflicts that arise from differing priorities. This analysis also highlights the challenges in reconciling economic development objectives with environmental sustainability and preserving traditional livelihoods. The following subsections provide detailed insights into each actor's role, objectives, and the implications of their activities for water resource management in the Lower Omo-Gibe River Basin.

2.2.1. Ethiopian Electric Power

Ethiopian Electric Power (EEP) is a state-owned enterprise responsible for generating, transmitting, and distributing electric power in Ethiopia. EEP's primary objective in the Lower Omo-Gibe Basin is to maximize hydropower generation from major projects such as the Gibe III and Koysha dams (Avery, 2017; Merrick, 2018). These projects are critical to Ethiopia's ambitious plans to become a regional energy hub and drive economic development (Zaniolo et al., 2021a).

The Gibe III Dam, completed in 2016, has a capacity of 1,870 MW, making it one of the largest hydroelectric projects in Africa (Avery, 2017). The Koysha Dam, part of the larger Omo-Gibe cascade,

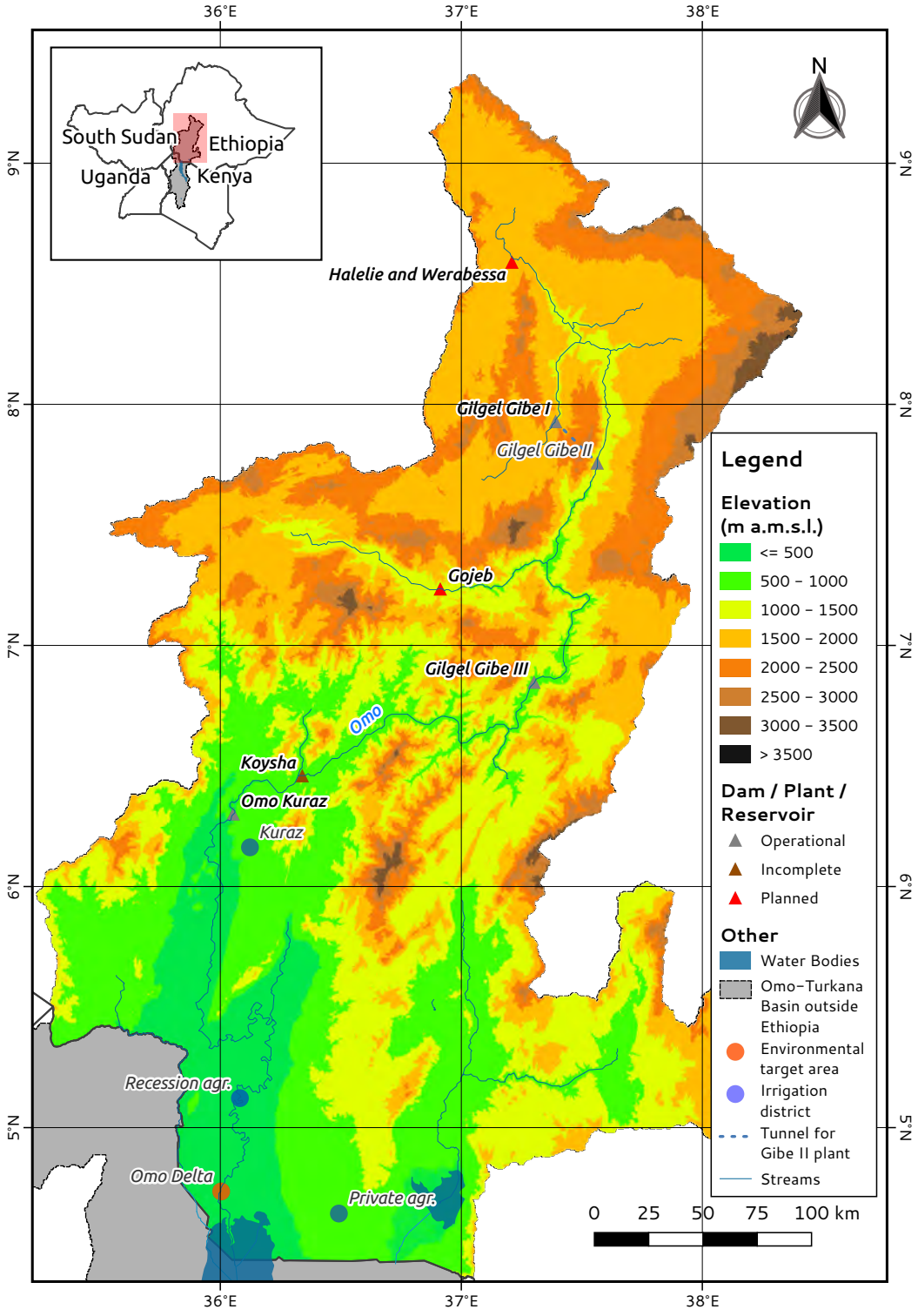


Figure 2.2: Digital elevation map of the Omo-Turkana Basin in Ethiopia, highlighting major hydrological and infrastructural components. This map reveals the basin's topographic diversity, ranging from high-altitude areas (brown and orange) to lowlands (green). Key features include major dams (triangles) and irrigation districts (circles), illustrating the extent of water resource development in the region. (Data sources: Hydrobasin boundary and component locations from DAFNE project (Micotti, 2020); elevation data from NASA Shuttle Radar Topography Mission (SRTM GL1) Global 30m database on OpenTopography.)

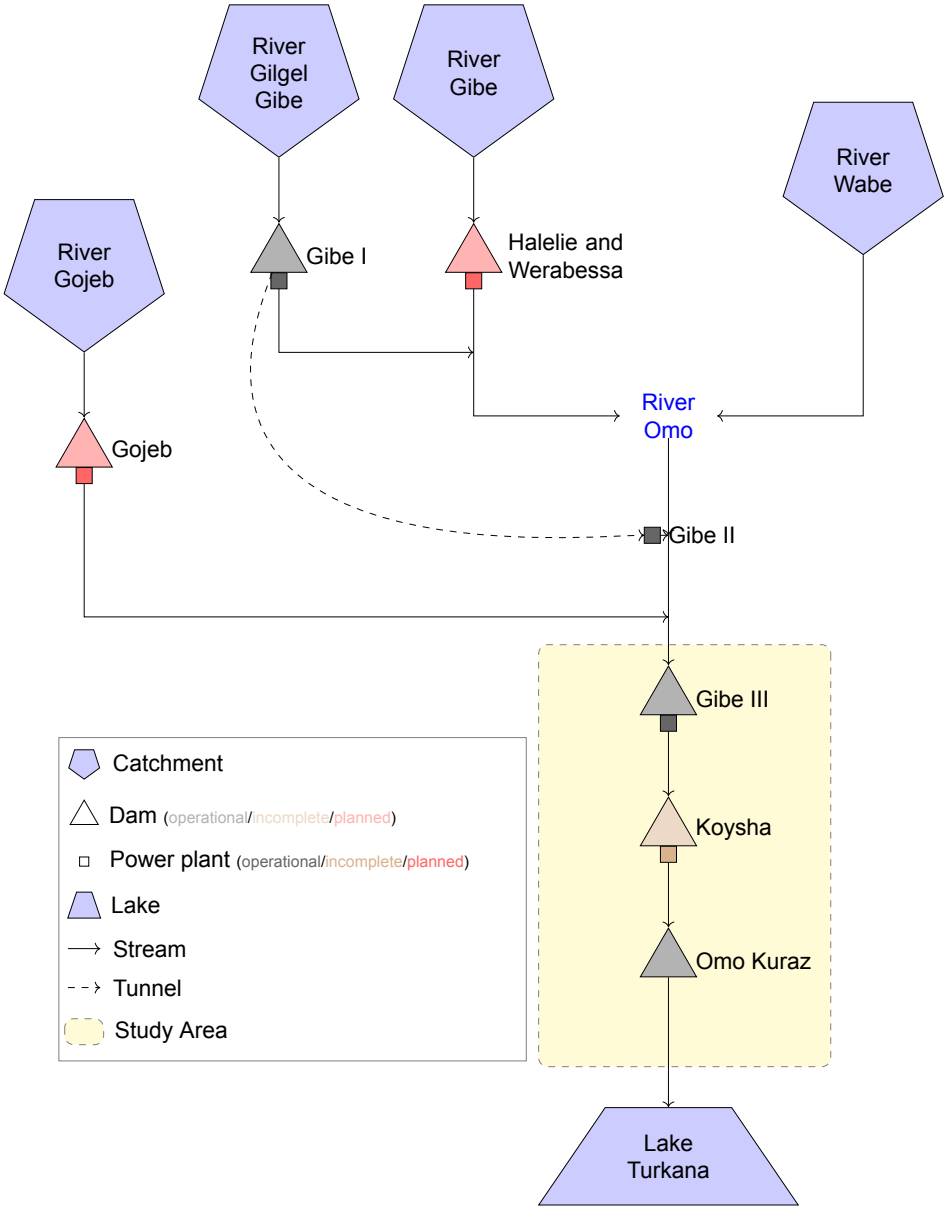


Figure 2.3: Schematic representation of the main components of the Omo River Basin. Key features include the major tributaries (Gilgel Gibe, Gibe, Gojeb, and Wabe rivers), existing and planned dams (e.g., Gibe I, II, III, Koysha, Gojeb, and Halelie & Werabessa), and associated power plants. This study focuses on the Lower Omo Basin (highlighted).

is expected to add another 2,160 MW of generating capacity upon completion (Zaniolo et al., 2021a). These projects align with Ethiopia's broader economic development goals, including increasing access to electricity, boosting industrial growth, and reducing poverty (Amos et al., 2021).

The hydropower generated by these dams is intended for domestic consumption and export to neighboring countries, potentially generating significant foreign exchange revenue for Ethiopia (Avery and Tebbs, 2018). However, this focus on maximizing hydropower generation has led to significant ecological and social trade-offs, including the alteration of natural flow regimes and impacts on downstream communities (Hodbod et al., 2019; Gebre et al., 2023).

EEP's perspective in water resource management is primarily driven by the need to ensure consistent and high water levels in the reservoirs to maintain optimal power generation. This often conflicts with other water uses and environmental concerns, highlighting the need for a balanced approach to water resource management in the basin (Zaniolo et al., 2021c).

2.2.2. Ethiopian Sugar Industry Group

The Ethiopian Sugar Industry Group (ESIG) is a semi-public organization tasked with managing and expanding Ethiopia's sugar industry. In the Lower Omo Basin, ESIG oversees the Omo Kuraz Sugar Development Project, one of the largest agricultural initiatives in the region (Avery, 2017; Gebeyehu et al., 2023). This project aims to transform the area into a major sugar production hub, leveraging the fertile floodplains of the Omo River for large-scale sugarcane cultivation.

ESIG's primary objective is to maximize sugar production to meet domestic demand and export targets, thereby contributing to Ethiopia's economic growth and food security (Amos et al., 2021). The project is expected to create thousands of jobs and stimulate local economies, aligning with national development goals (Hodbod et al., 2019).

However, developing vast sugar plantations has required substantial land conversion and water diversion from the Omo River. This has led to concerns about environmental sustainability and the displacement of indigenous communities (Hodbod et al., 2019; Amos et al., 2021). The water-intensive nature of sugarcane cultivation poses challenges for maintaining adequate river flows, which are essential for downstream ecosystems and traditional livelihoods (Avery and Tebbs, 2018; Zaniolo et al., 2021c).

ESIG's perspective in water resource management is centered on ensuring a reliable and abundant water supply for irrigation throughout the year. This need often competes with other water uses, particularly during dry seasons, highlighting the importance of efficient water allocation strategies (Gebre et al., 2023).

2.2.3. Local Communities

The local communities in the Lower Omo-Gibe River Basin comprise a diverse group of agro-pastoralists, fishermen, and small-scale farmers who have relied on the river's natural flow regime for generations (Hodbod et al., 2019; Avery, 2017). These communities practice traditional livelihood strategies such as recession agriculture, where crops are planted on the floodplains as the river recedes after seasonal floods, providing fertile soil and moisture for cultivation (Carr, 2017).

Fishing is another critical source of food and income for many local communities, with many relying on the fish stocks of the Omo River and Lake Turkana (Avery and Tebbs, 2018). The seasonal flooding of the Omo River has historically played a crucial role in sustaining these livelihoods and shaping the cultural practices of the local communities (Amos et al., 2021).

However, the construction of the Gibe III and Koysha dams has significantly disrupted these traditional practices by altering the timing and magnitude of floods, reducing sediment deposition, and changing fish migration patterns (Avery and Tebbs, 2018; Amos et al., 2021). These changes have led to increased food insecurity and economic hardship for many local communities (Hodbod et al., 2019).

Local communities advocate for ensuring environmental flows in the Omo-Gibe River Basin to sustain their agricultural activities and preserve their cultural heritage. They have raised concerns about the long-term viability of their livelihoods in the face of large-scale industrial and hydropower developments, calling for more inclusive and sustainable water management practices (Hodbod et al., 2019; Gebre et al., 2023).

The perspective of local communities in water resource management is focused on maintaining sufficient river flows to support traditional agricultural practices and fishing activities. This often conflicts

with the water demands of large-scale hydropower and irrigation projects, highlighting the need for a balanced approach that considers the needs of all stakeholders (Zaniolo et al., 2021c).

2.3. Literature Review

Hydrological modeling in the Lower Omo Basin has been a significant area of research due to the environmental, social, and economic implications of major infrastructure projects like the Gibe III and Koysha dams. This literature review synthesizes findings from various studies, highlighting the hydrological changes, ecological impacts, socio-economic consequences, broader regional implications, and the applications of advanced modeling techniques in water resource management.

2.3.1. Hydrological Modeling and Climate Change Impacts

The construction of the Gibe III Dam has significantly altered the flow regime of the Omo River, reducing seasonal variability crucial for downstream ecosystems and agricultural practices. Chaemiso et al. (2016) utilized the SWAT model to simulate hydrological responses in the Omo-Gibe river basin, indicating significant variations in future precipitation and temperature patterns. Their study showed an overall increase in annual temperature and variations in monthly and seasonal precipitation from the base period 1985-2005.

Building on this work, Dagne et al. (2023) employed remote sensing techniques and hydrological modeling to assess the spatiotemporal dynamics of water resources in the basin. Their findings highlighted the importance of considering climate and land use changes in future water resource management strategies.

Lukas et al. (2023) developed a machine learning-based approach for predicting streamflow in data-scarce regions, using the Omo-Gibe Basin as a case study. Their model demonstrated high accuracy in predicting streamflow under various climate scenarios, providing a valuable tool for water resource planning in the region.

Moreover, Dondeyne et al. (2021) explored the intersections of climate change, large-scale land investments, and local livelihoods in the Lower Omo Valley. Their research emphasized the importance of considering multiple stressors in vulnerability assessments and adaptation planning.

The importance of addressing climate change impacts on water resources is further emphasized by Voinov et al. (2016), who highlight the need for more stakeholder engagement and adaptive management approaches in modeling efforts. This aligns with the seminal work of Milly et al. (2008), who argued that stationarity can no longer serve as a central assumption in water-resource risk assessment and planning, emphasizing the need for new approaches to water management in the face of climate change.

2.3.2. Ecological Impacts and Environmental Flows

Research has extensively explored the ecological impacts of hydrological changes in the Omo-Gibe Basin. Hodbod et al. (2019) investigated the social-ecological changes in the Omo-Turkana basin, focusing on the impacts of altered hydrology on local ecosystems. The regulation of river flow by the Gibe III Dam has reduced the natural flood pulses essential for maintaining the ecological balance of the region.

Avery (2017) highlighted the potential threats to fish populations and other aquatic organisms that depend on seasonal flooding for spawning and feeding. Their work emphasized the need for maintaining environmental flows to support ecosystem services and biodiversity in the basin.

Niggemann and Graf (2020) studied the impact of dam construction on sediment transport and turbidity in the Omo River. Their findings revealed significant reductions in sediment load downstream of the dams, with potential long-term consequences for the river's geomorphology and aquatic ecosystems.

Further, Carr (2017) provided a comprehensive analysis of the impacts of river basin development on human rights in Eastern Africa, including the Omo-Gibe Basin. Their work underscored the importance of considering social and environmental justice in water resource management decisions.

2.3.3. Socio-economic Consequences and Stakeholder Perspectives

The socio-economic impacts of the Gibe III and Koysha dams are profound and multifaceted. Amos et al. (2021) documented the challenges pastoralist communities face due to altered flooding patterns,

which have disrupted traditional flood-retreat agriculture. Additionally, the construction of these dams has necessitated the resettlement of local communities, often without adequate compensation or support, leading to significant socio-economic hardships (Hodbod et al., 2019).

Hailu Woldegebrael (2018) investigated the process of land acquisitions for large-scale agriculture in the Lower Omo Valley. Their work revealed the complex power dynamics and conflicts of interest involved in the transformation of the basin's landscape.

The broader regional implications of the Gibe III and Koysha dams extend beyond Ethiopia. Zaniolo et al. (2021c) discussed the broader implications of these projects, emphasizing the need for comprehensive environmental impact assessments and strategies to mitigate adverse effects. Moreover, the impacts extend beyond Ethiopia, affecting Lake Turkana in Kenya, which relies on the Omo River for its inflow (Avery and Tebbs, 2018).

2.3.4. Advanced Modeling Techniques in Water Resource Management

In recent years, advanced modeling and optimization techniques have been applied to address the complex challenges of water resource management in the Omo-Gibe Basin and similar contexts.

The Evolutionary Multi-Objective Direct Policy Search (EMODPS) framework has emerged as a powerful tool in water resource management, addressing complex multi-objective decision-making problems. Giuliani et al. (2014) employed EMODPS to identify and refine reservoir policies that reduce policy inertia and myopia in water management. Their approach provided a comprehensive evaluation of the trade-offs in managing large-scale water systems, highlighting the flexibility and robustness of EMODPS.

Quinn et al. (2018) applied EMODPS to investigate the impact of changing monsoonal dynamics and human pressures on multi-reservoir management. Their study revealed the capability of EMODPS to handle the complexities of evolving hydrological patterns and stakeholder needs, providing a nuanced understanding of the interactions between natural and human systems.

Zaniolo et al. (2021b) focused on the Omo-Gibe River Basin, using EMODPS to design multi-objective reservoir policies that balance hydropower production with ecological and social considerations. This application underscored the importance of incorporating diverse objectives and stakeholder perspectives in water resource management.

One of the key strengths of EMODPS is its ability to manage deep uncertainty in water resource systems. Giuliani et al. (2016) highlighted the framework's effectiveness in optimizing reservoir operations under uncertain future conditions, such as climate change and socio-economic developments. By exploring a wide range of potential futures, EMODPS helps decision-makers identify robust policies that perform well across different scenarios.

Herman and Giuliani (2018) further demonstrated the utility of EMODPS in addressing non-stationary and dynamic systems. Their research showed how EMODPS could adaptively manage water resources by continuously updating policies based on new information, thus enhancing the resilience of water management strategies.

Building on these advances, Zaniolo et al. (2021b) contributed a novel method to learn the optimal policy representation by combining feature selection with multi-objective Direct Policy Search. This approach allows for dynamic adaptation of the policy input set based on the objective trade-off, further enhancing the flexibility and performance of water management strategies.

The performance of evolutionary algorithms in multi-objective water resource management has been extensively studied. Zatarain Salazar et al. (2016) provided a comprehensive diagnostic assessment of modern MOEAs' abilities to support EMODPS using the Conowingo reservoir in the Lower Susquehanna River Basin as a case study. Their findings highlighted the challenges and considerations in applying EMODPS to water resource problems and identified promising algorithmic approaches.

2.3.5. Challenges in Hydrological Modeling

Despite the advances in hydrological modeling and optimization techniques, significant challenges remain in model transparency, reproducibility, and extensibility. Hutton et al. (2016) highlighted that most computational hydrology is not reproducible, questioning its scientific validity. They emphasized the need for improved documentation, version control, and open-source practices in hydrological modeling.

Jakeman et al. (2006) proposed ten iterative steps for developing and evaluating environmental models, emphasizing the importance of transparency in model conceptualization, implementation, and

evaluation. Their work highlighted the need for clear communication of model assumptions, limitations, and uncertainties.

The issue of model extensibility is particularly relevant in the context of evolving water management challenges. Pahl-Wostl (2007) discussed the need for adaptive management approaches in water resources, emphasizing the importance of flexible and extensible modeling frameworks that can incorporate new data, scenarios, and management objectives as they emerge.

Examining existing hydrological modeling repositories reveals that many are tailored to specific contexts and need more modularity and flexibility for easy adaptation to new scenarios. For instance, the CALFEWS model by Zeff et al. (2021), while comprehensive for California's water-energy system, is particular to that region and would require significant modifications for application in other contexts.

Similarly, while innovative in its approach, the Lake Problem Direct Policy Search model by Quinn et al. (2017) is tailored to a specific lake system and needs more generalizability for application to diverse water resource management scenarios.

These examples underscore the need for more modular, extensible, and general-purpose modeling frameworks in water resource management. Such frameworks would enhance the reproducibility of research, facilitate the broader application of advanced optimization techniques like EMODPS, and enable more effective knowledge transfer between different water management contexts.

2.3.6. Research Gap

Despite extensive research on the hydrological and socio-economic impacts of the Gibe III and Koysha dams, there remains a significant gap in the transparent and efficient modeling of the water resource system in the Lower Omo Basin. This gap is particularly evident in the context of competing water uses and the uncertainties arising from evolving development projects and data scarcity.

Current implementations of water resource models often involve custom simulation and optimization code with hardcoded relationships, which significantly affects the reproducibility, transparency, and extensibility of the models. As highlighted by Voinov et al. (2016), there is a pressing need for modeling approaches that are more accessible to a broader range of stakeholders and can adapt to changing management objectives and scenarios.

Moreover, there is a notable lack of a general-purpose modeling framework that integrates the Evolutionary Multi-Objective Direct Policy Search (EMODPS) for water resource systems. Such a framework would enhance the reproducibility of research and facilitate the broader application of EMODPS in various water management contexts.

This research aims to address these gaps by creating a transparent and reproducible framework that integrates EMODPS, thereby advancing the state-of-the-art in water resource modeling and management. We will also investigate the trade-offs among water uses in the Lower Omo-Gibe River Basin, by searching for Pareto-approximate policy choices using the developed framework. This approach aligns with the call for more integrative and adaptive water resource management strategies, as advocated by Biswas (2004) and others in the field.

Furthermore, the global context of water resource challenges, as highlighted by Vörösmarty et al. (2000), underscores the importance of developing robust, adaptable modeling frameworks that can address the complex interplay of climate change, population growth, and competing water demands. By addressing these challenges in the context of the Lower Omo-Gibe River Basin, this research contributes to the broader effort of improving water resource management in the face of global environmental change.

3

Hydrological Modelling of the Lower Omo-Gibe Basin

This chapter presents a comprehensive overview of the hydrological modeling approach employed for the Lower Omo-Gibe Basin. We begin by delineating the basin structure, followed by a detailed description of the data requirements and sources. The data includes flow rates, evaporation rates, reservoir properties, irrigation schedules, and environmental flow requirements. The chapter then elaborates on the modeling approach, including key assumptions and limitations.

3.1. Basin Structure

The Lower Omo-Gibe Basin is a complex hydrological system characterized by multiple tributaries, reservoirs, and irrigation schemes. Our study focuses on the lower part of the basin, encompassing the Gibe III Dam, the upcoming Koysha Dam, and the extensive Omo-Kuraz sugar plantations. Figure 3.1 presents a schematic representation of the basin as conceptualized in our model. The The next chapter will discuss the theoretical framework employed for water resource modeling.

While several tributaries exist in the lower basin, the primary inflow into the **Gibe-III** Dam comes from the Gibe River. We call this **Flow A** in our model. The release from the Gibe-III Dam, labeled as **Gibe-III Release**, combines with the inflow from other tributaries that we collectively refer to as **Flow B** before reaching the Koysha Dam. The **Koysha** Dam is a new addition to the basin and is expected to impact the water flow and management in the region significantly. The release from the Koysha Dam goes to a major weir labeled **Koysha Headworks** that diverts water for the Omo-Kuraz sugar plantations. The water diverted for irrigation is labeled by **Canals to Kuraz Sugar Plantations** in the model. The remaining water flows downstream through the existing **Main Channel**. The contribution of the rest of the catchment from Koysha to the town of **Omorate** is modeled as **Flow C** in the model. A significant part of Flow C is the Usno tributary that forms from the confluence of the Magi and Neri rivers. The merged flow from the main channel and Flow C forms the **Outflow to Lake Turkana**. In addition to releases from the dam, water is also lost through evaporation from the reservoirs, which is modeled as **Gibe-III Evaporation** and **Koysha Evaporation**.

The irrigation demand in our model is to be met by the water diverted from the Kuraz Headworks to Kuraz Sugar Plantations. While multiple physical canals are used to distribute water to the plantations, we have simplified the model by considering a single canal representing the total water diverted for irrigation. No major cities or towns are in the lower basin, and the water demand for domestic and industrial use is assumed to be negligible. Most of the domestic (primarily from those working on the sugar plantations) and industrial (primarily sugar mills) water demand is related to the sugar plantations and is accounted for in the irrigation demand.

The environmental flow requirement is the minimum amount of water that must be maintained in the river to sustain the ecological balance of the basin. In our model, we take the water flow at the Omorate town as the reference point for the environmental flow requirements. The water flow at Omorate is a sum of the main channel flow and the Flow C. The environmental flow requirements are modeled as the minimum flow that must be maintained at Omorate to sustain the ecological balance of the basin.

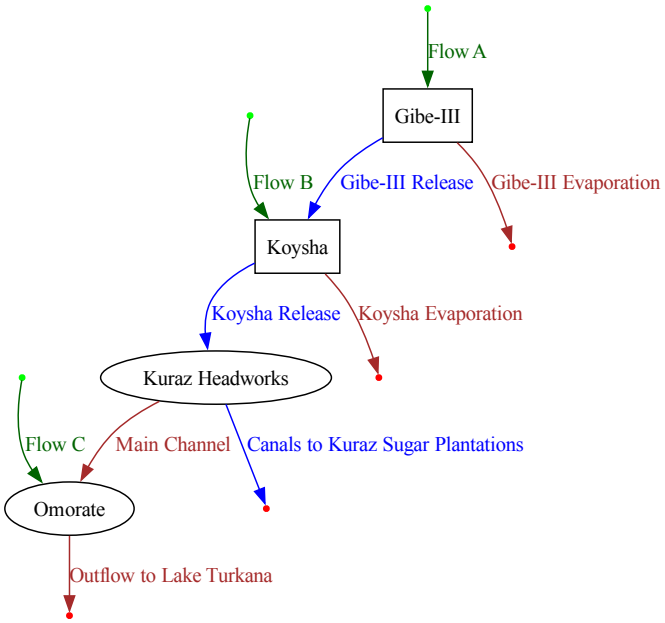


Figure 3.1: Schematic representation of the Lower Omo-Gibe Basin. This diagram illustrates the key hydrological components and their interconnections. Flow A represents the primary inflow into the Gibe-III Dam from the Gibe River. The Gibe-III Release combines with Flow B before reaching the Koysha Dam. Water is then diverted at the Koysha Headworks for irrigation (Canals to Kuraz Sugar Plantations), with the remainder continuing through the Main Channel. Flow C represents the additional catchment contribution between Koysha and Omorate. Evaporation losses from both reservoirs are explicitly modeled.

The environmental flow contributes not only to the aquatic vegetation and wildlife in the Omo Delta and Lake Turkana but is also relevant to the traditional livelihoods of the local communities dependent on the river for fishing and recession agriculture.

3.2. Data Requirements

The hydrological model requires several data inputs to simulate the water flow and management in the basin. The primary data requirements include:

- Flow rates for external inflows
- Reservoir properties such as storage capacity, bathymetry, and operational constraints
- Characteristics of the hydropower plants associated with the dams
- Evaporation rates from the reservoirs
- Irrigation schedules and resulting water demand
- Environmental flow requirements
- Infrastructure commissioning and operational timelines

3.3. Data Description

In this section, we discuss the data available for each of these requirements and how they are used for the hydrological modeling of the Lower Omo-Gibe Basin in our study.

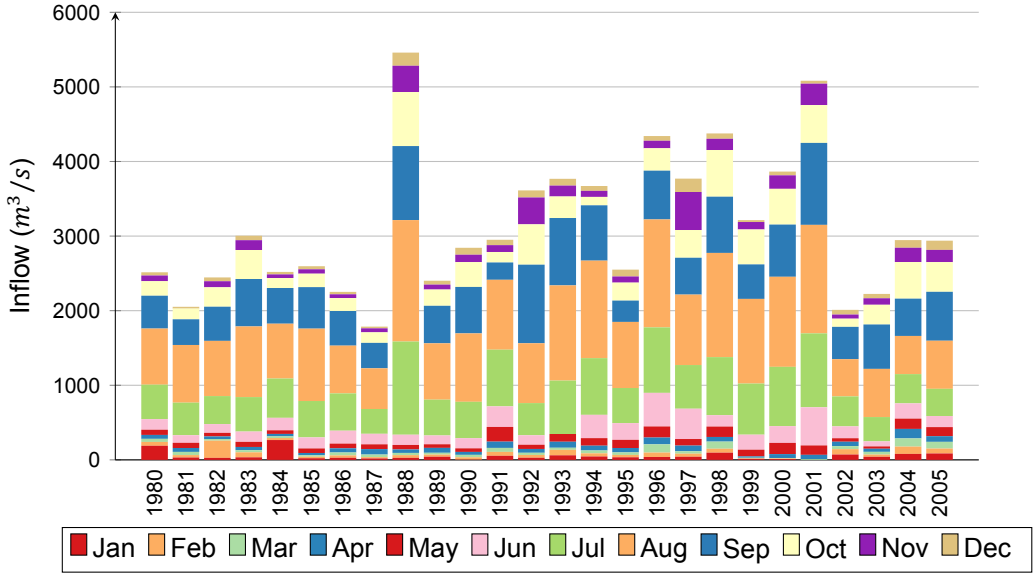


Figure 3.2: Annual inflow to Gibe III during 1980-2005. This stacked bar chart depicts the monthly contributions to annual inflow over 26 years. Each color represents a different month, allowing for visualization of seasonal patterns and inter-annual variability. The significant year-to-year fluctuations in total inflow and the changing proportions of monthly contributions highlight the hydrological complexity of the basin.

3.3.1. Flow Rates

Our hydrological model needs flow rates for Flow A, B, and C. The historical rates were available only for inflow to the Gibe-III Dam, which we use as a proxy for Flow A. Figure 3.2 presents the annual inflow patterns to Gibe-III from 1980 to 2005, illustrating significant inter-annual variability. Figure 3.3 provides a more detailed view of the seasonal inflow patterns, presenting the average monthly inflows as box plots. It shows the 25th and 75th percentiles, the median, and the outliers for each month. Historical records indicate extreme variability in inflow rates, with dramatic differences observed from year to year, particularly in the critical month of August. For example, in August 1988, the mean monthly inflow was extremely high at 1628 m^3/s, contrasting sharply with a low of 499 m^3/s in August 2002.

Our model uses the mean monthly inflow rates to the Gibe-III Dam as a proxy for Flow A. However, historical data for the inflow rates to the Koysha Dam and the Omorate town could not be obtained. We therefore estimate the inflow rates for Flow B and Flow C based on the catchment area and the historical data for the Gibe-III Dam. The average annual outflow of Omo River to Lake Turkana has been estimated as 438 m^3/s by International Rivers (2009). Since the outflow to Lake Turkana is the sum of the three external flows (Flow A, Flow B, and Flow C) in our model and we the average annual rate for Flow A, we can estimate the inflow rates to the Koysha Dam and the Omorate town based on the historical data for the Gibe-III Dam and the share of the catchment area between the Gibe-III Dam and the Koysha Dam and between the Koysha Dam and the Omorate town. Here, we assume that the precipitation and evaporation rates are uniform across the basin, and the inflow rates are proportional to the catchment area. Using this approach, we estimate the monthly inflow rates for Flow B and Flow C as 40% and 28% of values for Flow A, respectively.

3.3.2. Reservoir Properties

The Gibe-III and Koysha Dams are the two primary reservoirs in the Lower Omo-Gibe Basin.

Table 3.1 summarizes the key properties of the Gibe-III and Koysha reservoirs.

The bathymetry data for the Gibe-III was available while the data for the Koysha Dam could not be obtained. We create synthetic bathymetry data for the Koysha Dam by comparing the two reservoirs' maximum volume, surface area, and head. Figure 3.4 illustrates the bathymetry data for both reservoirs, showing the relationships between head, surface area, and volume.

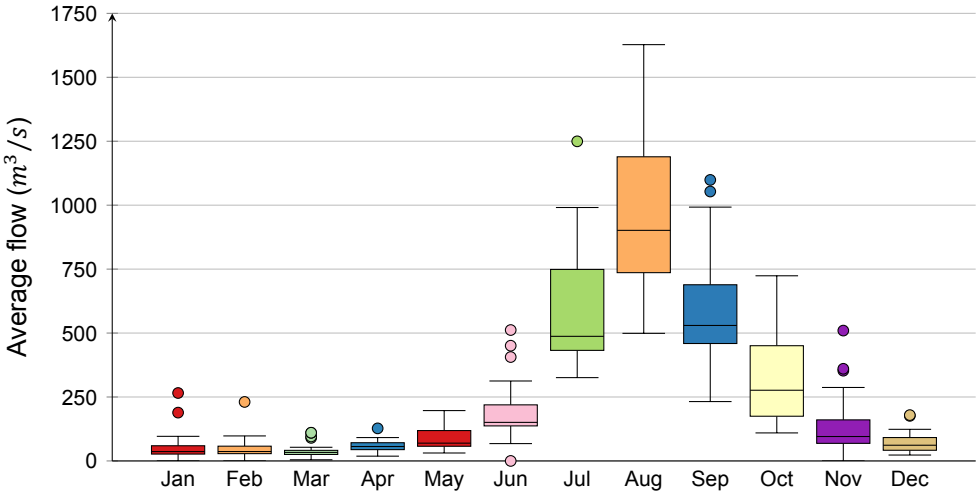


Figure 3.3: Average monthly inflow to Gibe III during 1980-2005. This box plot visualizes the statistical distribution of inflows for each month. The boxes represent the interquartile range (IQR), with the median as a horizontal line. Whiskers extend to the highest and lowest values within 1.5 times the IQR. Outliers, if any, are plotted as individual points. This representation clearly illustrates the seasonal pattern of inflows, with peak flows typically occurring in August and the lowest flows in March.

Table 3.1: Reservoir Properties for Gibe-III and Koysha Dams. This table presents key physical and operational parameters for the two main reservoirs in the Lower Omo-Gibe Basin. MCM denotes million cubic meters. The operational volumes indicate the range within which the reservoirs are (assumed to be) typically managed.

	Gibe-III	**Koysha**
Maximum Volume (MCM)	14,000	6,000
Maximum Surface Area (km²)	200	119
Maximum Head (m)	220	178.5
Maximum Operational Volume (MCM)	11,750	6,000
Minimum Operational Volume (MCM)	7,000	3,000

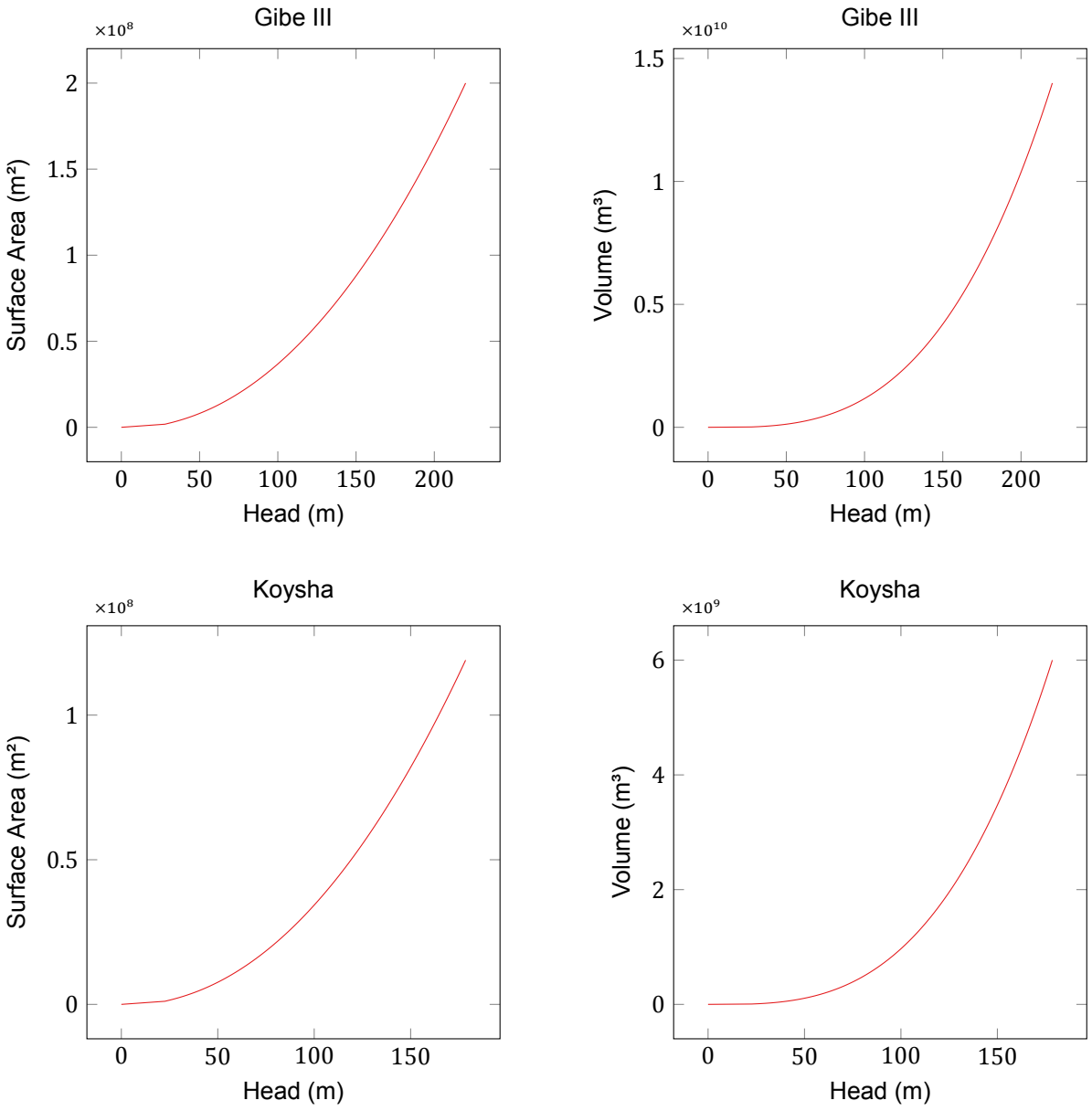


Figure 3.4: Surface Area and Volume vs Head for Gibe III and Koysha reservoirs. These plots illustrate the relationships between both reservoirs' water level (head), surface area, and volume. These relationships' non-linear nature is crucial for accurately modeling reservoir dynamics, including evaporation losses and hydropower generation potential at different storage levels. Note that the Koysha data is synthetic, derived from comparisons with Gibe III due to a lack of direct measurements.

Table 3.2: Hydropower Plant Characteristics for Gibe-III and Koysha Dams. This table provides essential technical parameters for the hydropower facilities. The installed capacity represents the maximum power output, while the generation capacity indicates the expected annual energy production. Turbine efficiency, head, and maximum flow are crucial for calculating power generation under various operational conditions.

	Gibe-III	**Koysha**
Installed Capacity (MW)	1870	2160
Generation Capacity (GWh/year)	6500	6460
Turbine Efficiency	0.90	0.90
Turbine Head (m)	9.0	8.5
Turbine Max Flow (m^3/s)	1064	1440

3.3.3. Characteristics of Hydropower Plants

The Gibe-III and Koysha Dams are both hydropower plants with significant generation capacities. There is some ambiguity regarding the plant capacities and generation rates, with different sources providing slightly different values (Webuild Group, 2024a,b; Studio Pietrangeli, 2024b,a). Table 3.2 presents the key characteristics of the hydropower plants associated with the Gibe-III and Koysha dams.

3.3.4. Evaporation rate for reservoirs

The monthly evaporation values are sourced from the WRF model to estimate the evaporation rates for the Gibe-III and Koysha reservoirs owing to their enhanced accuracy over the older MOD16 dataset (Mathewos et al., 2022). Figure 3.5 illustrates the monthly evaporation rates used in our model.

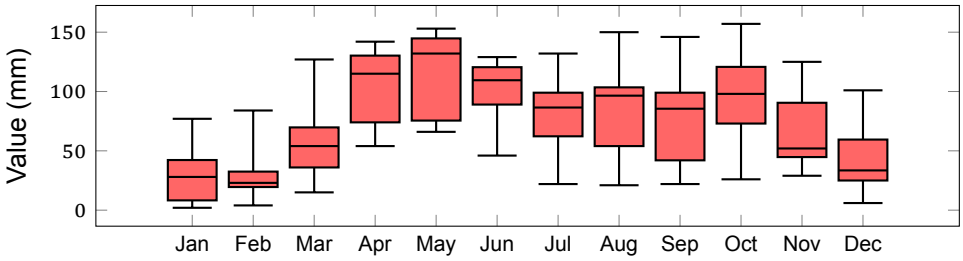


Figure 3.5: Monthly evaporation rates derived from the WRF model. This box plot illustrates the statistical distribution of evaporation rates for each month. The boxes represent the interquartile range (IQR), with the median as a horizontal line. Whiskers extend to the highest and lowest values within 1.5 times the IQR, and outliers are plotted as individual points. The plot reveals a clear seasonal pattern, with peak evaporation typically occurring in the warmer months (April to October) and lower rates in the cooler months (November to March).

The WRF model estimates a mean monthly evaporation rate of 72.79 mm with a standard deviation of 26.58 mm for the region. These values are crucial for accurately modeling reservoir water losses, which can significantly impact water availability for downstream uses.

3.3.5. Irrigation Demand in Lower Omo Valley

The Omo-Gibe basin's potential for irrigation has been well recognized, with estimations suggesting the total irrigable area could be anywhere from 73,000 hectares to an ambitious 445,000 hectares. A more focused Environmental and Social Impact Assessment (ESIA) identified 142,000 hectares as particularly suitable for large-scale irrigation projects in the lower basin.

In 2011, the state-owned Ethiopian Sugar Corporation (ESC) launched the Kuraz Sugar Development Project (KSDP) aims to expand sugarcane cultivation in the Omo Valley significantly. The project's initial plan envisioned developing estates spanning 175,000 hectares, supplemented by the construction of five processing factories.

However, the scale of the project has undergone considerable fluctuations. The actual pace of expansion has not kept up with initial timelines. By October 2014, only about 10,000 hectares of Block I had been cleared for cultivation on the left side of the river, with 6,600 hectares planted with sugarcane. Although planting in Kuraz I began in 2014 with an anticipated production start in 2016, and Kuraz II and III were expected to start production by 2017, these targets were not met. In fact, by the end

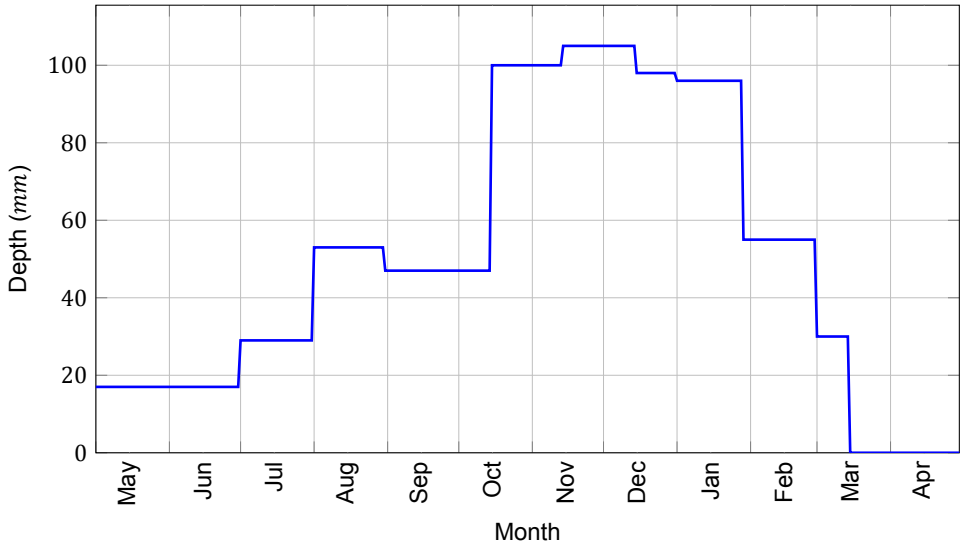


Figure 3.7: Irrigation Schedule for Omo-Kuraz Sugar Plantations. This graph shows the region's monthly irrigation water requirements for sugarcane cultivation. The y-axis represents the water depth needed in millimeters, while the x-axis shows the months of the year. The schedule reveals a distinct seasonal pattern, with peak water demand occurring during the drier months (November to January) and lower demand during the wetter months (July to September).

of 2018, no sugar production had commenced in South Omo. Figure 3.6 illustrates the expansion of agricultural estates in the Omo-Kuraz region from 2010 to 2020. Theoretically, if the three operational processing factories (Kuraz I, II, and III) were to run at full capacity, utilizing their 12,000 tons crushed per day (tcd) each, around 75,000 hectares of land could be actively cultivated.

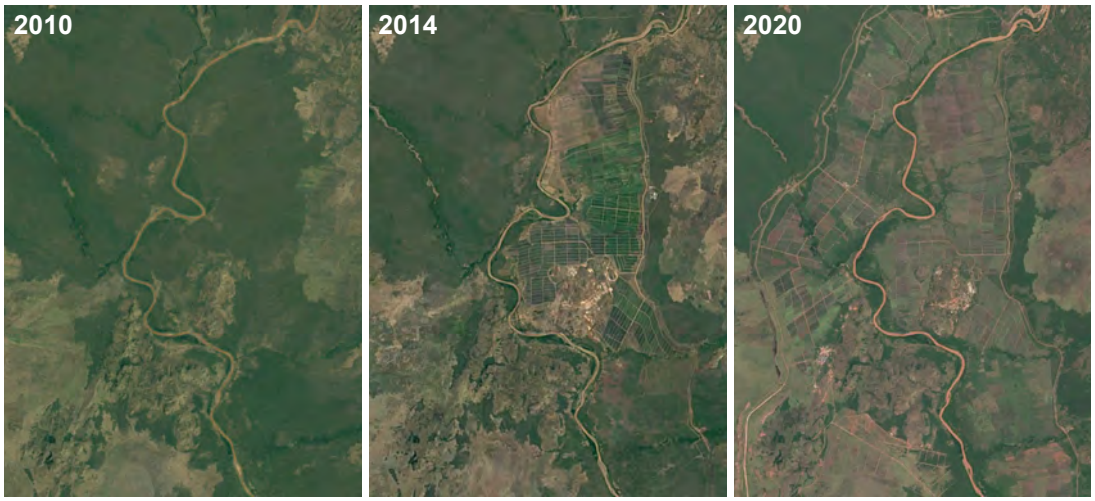


Figure 3.6: Expansion of agricultural estates in the Omo-Kuraz region over the years 2010, 2014, and 2020. These satellite images clearly illustrate the rapid transformation of the landscape from predominantly natural vegetation to large-scale agricultural development. The geometric patterns visible in the 2014 and 2020 images represent irrigated fields, indicating the significant increase in water demand for agriculture over this period.

The irrigation schedule for the Omo-Kuraz sugar plantations has been taken from the Gelalacha et al. (2024) specifying the region's water requirements for sugarcane cultivation. Figure 3.7 shows the irrigation schedule for the Omo-Kuraz sugar plantations.

3.3.6. Environmental Flow Requirements

Ethiopia does not have a formal environmental flow requirement for any of its rivers. We use 33% of the mean monthly flow at Omorate before the commissioning of the Gibe-III Dam as the environmental flow requirement for the Lower Omo-Gibe Basin. The average annual flow at Omorate is estimated as 438

m^3/s (International Rivers, 2009). The average annual environmental flow requirement is therefore estimated as 146 m^3/s. However, we use the monthly values for the environmental flow requirements in our model to account for the seasonal variations in the flow rates.

3.3.7. Infrastructure Commissioning and Operational Timelines

The Gibe-III Dam was commissioned in 2016. The construction of the Koysha Dam started in the same year and was expected to be completed by 2022. However, due to delays in construction, only 61% of the work has been completed by September 2023. Based on the current progress, we assume the Koysha Dam will be operational by 2029. We, therefore, simulate the Lower Omo-Gibe Basin from 2029 onwards in our model. We take a simulation horizon of 12 years, from 2029 to 2040, to analyze the tradeoffs between hydropower generation, irrigation demand, and environmental flow requirements in the basin.

3.4. Modelling Approach

For this project, we chose mechanistic models based on theories of physical system operations, focusing on mass-balance equations for reservoir management:

$$V_{t+1} = V_t + Q_{t+1} - E_{t+1} - R_{t+1}$$

Here, V_t represents initial reservoir volume, Q_{t+1} is inflow, E_{t+1} is net evaporation, and R_{t+1} is release volume.

We selected a monthly decision interval to balance seasonal changes and simulation feasibility. Calculating R_t adds significant computational load due to fresh release decisions and the need to adhere to physical and normative constraints, necessitating high time fidelity. To manage this, we integrated actual releases over 12-hour intervals.

3.5. Modeling Assumptions and Limitations

This section outlines the key assumptions and limitations of our Lower Omo-Gibe Basin model. We discuss system boundaries, basin structure simplifications, and steady-state versus dynamic components. These considerations contextualize our findings and underscore the need for an adaptable modeling framework, which is crucial for interpreting results and identifying areas for future refinement.

3.5.1. System Boundaries

The model presupposes that the hydrological and socio-economic systems of the basin can be effectively represented by quantifiable parameters such as water flow, reservoir capacities, and irrigation demand. It simplifies complex real-world interactions into measurable inputs and outputs to facilitate computational feasibility and clarity in policy implication analysis.

3.5.2. Simplifications in the Basin Structure

While there is some distance between the Koysha Dam and the Kuraz Headworks, resulting in some gains through the intermediate tributaries, for the sake of simplicity, we have simplified the model by assuming that the Kuraz Headworks receives the same amount of water as released from the Koysha Dam. We also don't account for groundwater recharge and discharge in the model or the time lags in the system.

3.5.3. Steady-State and Dynamic Components

It is assumed that specific physical characteristics, like basin topography and reservoir structures, remain static over the simulation period. In contrast, dynamic elements, such as water inflow and demand patterns, are modeled as time-varying, reflecting real-world fluctuations and uncertainties.

3.5.4. Implications for Overall Modeling Framework

The assumptions and limitations of the hydrological model above highlight the need for an extensible and adaptable modeling framework that can incorporate new data and insights as they become available. For instance, we can refine our model if we can obtain more detailed data on the flow rates in the tributaries between Koysha and Kuraz. Similarly, as the Omo-Kuraz sugar plantations expand and new

irrigation schemes are developed, we can update our model to reflect these changes. This requires a general-purpose modeling framework that can be easily extended and modified to incorporate new data and assumptions about the basin. The generalization can also allow us to apply the same framework to other basins, making it a versatile tool for water resource management. In the next chapter, we will present and discuss the HydroWizard framework that aims to address these requirements and demonstrate its application to the Lower Omo-Gibe Basin.

4

Methodology

This chapter presents the methodologies for modeling and optimizing water resource management in the Lower Omo-Gibe River Basin. It introduces three key components: the established XLRM Framework for complex system analysis and the Evolutionary Multi-Objective Direct Policy Search (EMODPS) for reservoir control optimization, followed by an in-depth exploration of the novel HydroWizard modeling framework. HydroWizard, developed for this study, offers a transparent, efficient, and reproducible approach to water resource modeling. The chapter details its architecture, functionality, and application to the Lower Omo-Gibe Basin, showcasing its innovative features such as the use of Directed Acyclic Graphs (DAGs) and advanced flow computation algorithms.

4.1. XLRM Framework: Analyzing Complex Systems

The XLRM framework, formalized by Lempert et al. (2003), offers a systematic method for organizing pertinent information in formal analysis, emphasizing a cyclical and iterative approach. This method entails continually revisiting and refining data throughout the analysis.

4.1.1. Components of XLRM Framework

The framework categorizes information into four distinct groups: Policy Levers (L), Exogenous Uncertainties (X), Measures (M), and Relationships (R).

Policy Levers encompass the immediate actions and strategies that decision-makers aim to investigate.

Exogenous Uncertainties refer to external factors beyond the decision-maker's control that could significantly influence the success of their strategies. These uncertainties are essential in scenario planning, as they identify the primary driving forces confronting decision-makers.

Measures are the criteria used by decision-makers and stakeholders to assess the desirability of various scenarios.

Relationships illustrate how different factors interact and impact each other over time based on the selected policy levers and the manifestation of uncertainties. These interactions are captured in scenario generator simulations, which model the potential evolution of future outcomes. Consequently, the XLRM framework functions as an intellectual bookkeeping tool, guiding the process of elicitation and discovery throughout the analysis.

4.1.2. XLRM Framework in the Lower Omo-Gibe Basin

The XLRM framework is particularly well-suited for analyzing the Lower Omo-Gibe River Basin's complex water resource management challenges. The framework's structured approach enables the identification of key policy levers, exogenous uncertainties, performance measures, and system relationships that influence the basin's water management strategies.

Figure 4.1 presents a comprehensive visualization of the XLRM framework applied to the Lower Omo-Gibe River Basin. This schematic elucidates the complex interplay between management decisions, external uncertainties, system relationships, and performance metrics, providing a robust foundation for our subsequent modeling and optimization efforts.

Policy Levers

The primary policy levers in the Lower Omo-Gibe Basin include the release decisions from Gibe III Dam, Koysha Dam, and the irrigation canals at Kuraz Headworks. These decisions directly impact hydropower generation, irrigation support, and environmental flow release in the basin. By optimizing these policy levers, decision-makers can balance the competing objectives of power generation, agricultural productivity, and ecological sustainability.

Exogenous Uncertainties

The basin's water management is influenced by various exogenous uncertainties, such as seasonal inflow variations, climate change impacts, and the expansion of the Omo-Kuraz sugar plantations.

Performance Measures

The performance of the water resource system in the Lower Omo-Gibe Basin is evaluated based on multiple criteria, including hydropower generation, irrigation support, and environmental flow release. These performance measures are benchmarks for assessing the effectiveness of different policy scenarios and guiding decision-making processes.

System Relationships

The interactions between the policy levers, exogenous uncertainties, and performance measures influence the overall system dynamics in the Lower Omo-Gibe Basin. By modeling these relationships and simulating different scenarios, decision-makers can gain insights into the trade-offs among competing objectives, enabling informed and strategic water management decisions. Chief among these relationships in the system is the mass balance relationship between inflows, outflows, and storage changes in the reservoirs.

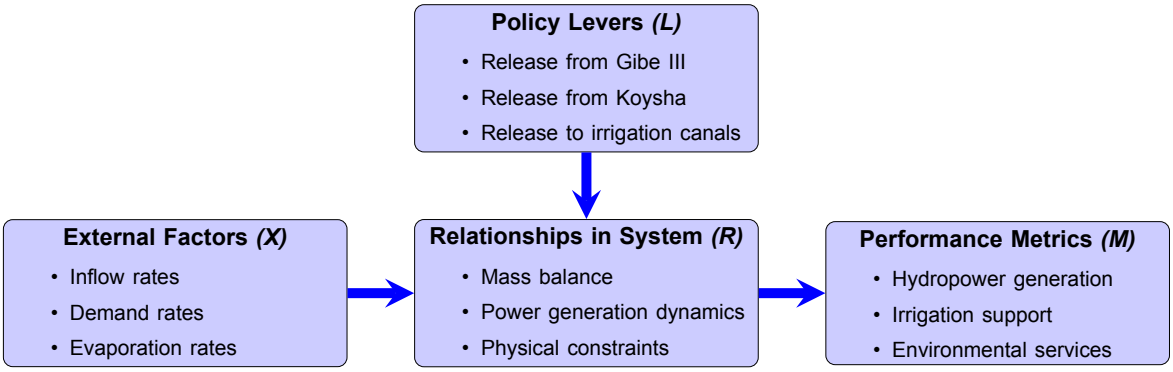


Figure 4.1: Schematic representation of the XLRM framework applied to the Lower Omo-Gibe River Basin. This diagram illustrates the interconnections between Policy Levers (L), Exogenous Uncertainties (X), Relationships in the System (R), and Performance Metrics (M).

4.2. EMODPS: Simulation-Optimization Framework for Reservoir Control

Evolutionary Multi-Objective Direct Policy Search (EMODPS), established by Giuliani et al. (2016), provides an advanced methodology for using Multi-Objective Evolutionary Algorithms (MOEAs) in the management of intricate multi-purpose reservoir systems. This method encompasses identifying reservoir policies, optimizing multiple objectives through evolutionary techniques and employing visual analytics to evaluate baseline operations and highlight critical operational trade-offs, aiding in the effective balancing of competing demands within a reservoir system.

Figure 4.2 provides a visual summary of the EMODPS methodology and illustrates how the various components of the system interact within the XLRM framework.

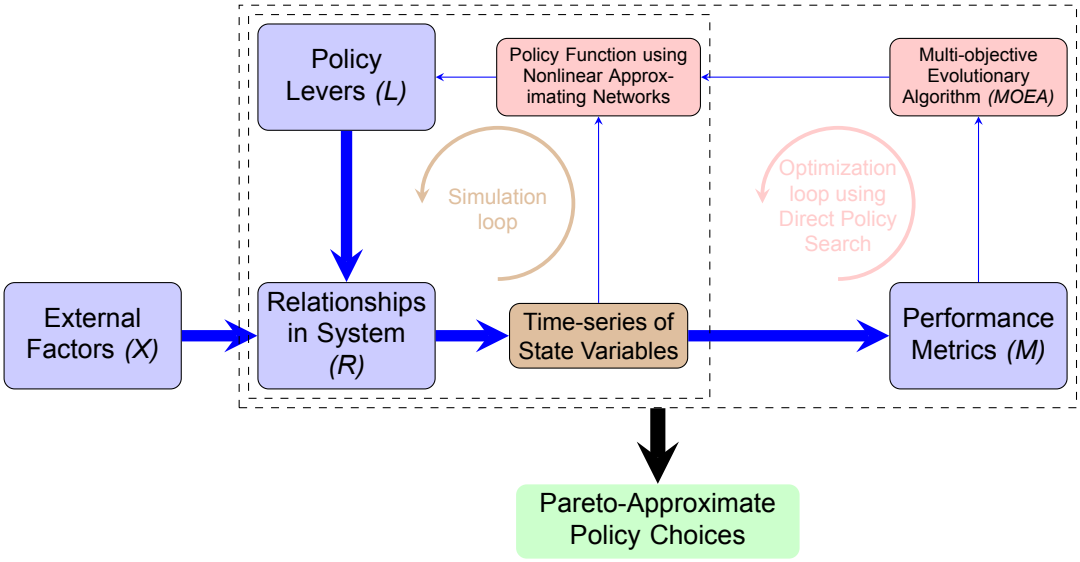


Figure 4.2: Integration of the Evolutionary Multi-Objective Direct Policy Search (EMODPS) framework within the XLRM structure for reservoir control optimization. This diagram illustrates the synthesis of EMODPS components (highlighted in red) with the XLRM framework (in blue), demonstrating the iterative process of policy function generation, simulation, and optimization leading to Pareto-approximate policy choices. The brown box represents the time series of state variables, a crucial link between system relationships and performance metrics that also serve as inputs to the policy function.

The EMODPS framework combines direct policy search with nonlinear approximating networks and multi-objective evolutionary optimization to determine Pareto-optimal control policies. The procedure involves two key phases: (i) Defining candidate operating policies through nonlinear approximators. (ii) Employing multi-objective evolutionary search to identify parameterizations that achieve Pareto-optimal trade-offs among conflicting management objectives.

EMODPS offers several notable advantages:

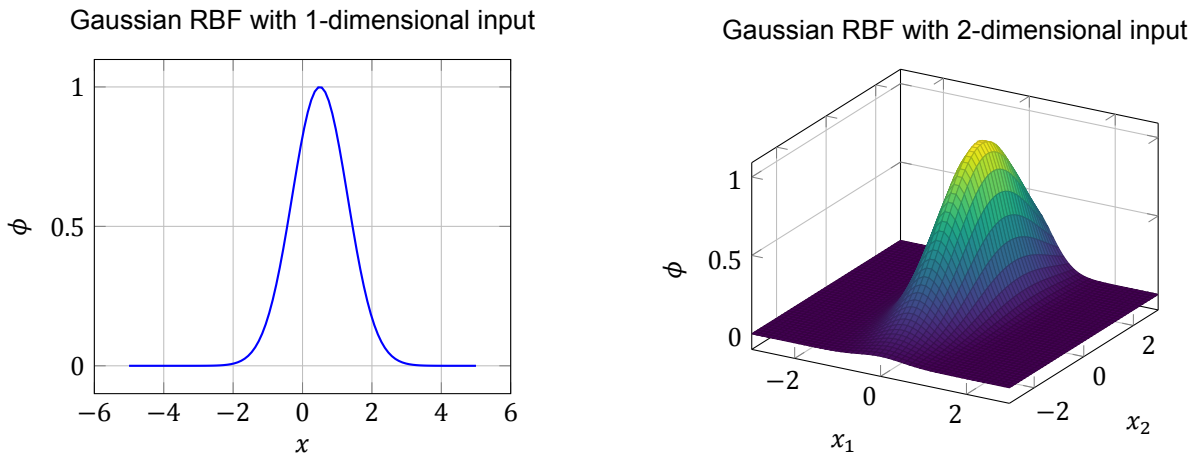
- It enables the simultaneous optimization of multiple objectives in a single run.
- It is not constrained by time-separable objective functions or model constraints.
- It is compatible with any simulation model, including those using Monte Carlo ensembles.
- It can incorporate external information into nonlinear approximators, enhancing the evaluation of uncertainties.

4.2.1. Components of EMODPS Framework

EMODPS framework has three primary components: Direct Policy Search (DPS), Radial Basis Function (RBF) Networks, and Multi-Objective Evolutionary Algorithms (MOEAs). Following is a brief overview of each component:

Direct Policy Search (DPS) optimizes decision-making policies directly from observed data and system states rather than generating a sequence of decisions. This method involves parameterizing control policies within a given family of functions (e.g., linear, piecewise linear, radial basis functions), simulating these policies, and then optimizing their parameters to perform best under simulation. DPS, also known as parameterization-simulation-optimization in the water resources literature, was first introduced by Rosenstein and Barto (2001) and further applied by Koutsoyiannis and Economou (2003). It handles complex, non-linear systems effectively, making it an advantageous alternative to traditional rule-based methods (Zatarain Salazar et al., 2016).

Compared to Stochastic Dynamic Programming (SDP), DPS offers several advantages. SDP, widely used for designing optimal reservoir operating policies, is challenged by the curse of dimensionality, the curse of modeling, and the curse of multiple objectives. In contrast, DPS can include multiple state variables in the optimized policies without requiring an explicit transition probability model, reducing computational complexity. DPS also allows the direct use of system simulations and can incorporate exogenous information into decision-making, enhancing performance under uncertain conditions.



(a) A 1D Gaussian RBF with $c = 0.5$, $\sigma = 0.8$.

(b) A 2D Gaussian RBF with $c = (0, 1)$, $\sigma = (0.5, 1.5)$.

Figure 4.3: Examples of Gaussian Radial Basis Functions (RBFs) for 1-dimensional and 2-dimensional inputs. These visualizations illustrate the non-linear mapping capabilities of RBFs, which are fundamental to their application in approximating complex control policies in water resource systems.

When combined with Multi-Objective Evolutionary Algorithms (MOEAs), DPS can efficiently optimize multiple objectives in a single run, making it a powerful tool for managing water resources and other applications requiring robust, flexible policy design (Giuliani et al., 2016).

A **radial basis function (RBF)** is a real-valued function, ϕ, that depends solely on the distance between the input and a fixed point. This fixed point can be the origin, yielding $\phi(x) = \hat{\phi}(\|x\|)$, or another specified point c, known as the center, resulting in $\phi(x) = \hat{\phi}(\|x - c\|)$. Several functions satisfy the radial basis function property, including the Gaussian function, the inverse multiquadric function, polyharmonic splines, and the thin-plate spline (Zatarain Salazar et al., 2024). RBFs are commonly used in interpolation, approximation, and machine learning applications due to their flexibility and ability to approximate complex functions with a small number of parameters (Buhmann, 2003).

Gaussian function, $\phi(x) = \exp\left(-\frac{\|x-c\|^2}{2\sigma^2}\right)$, where σ is the width parameter determining the function's spread, is an example of an infinitely smooth radial basis function. Figure 4.3 illustrates examples of Gaussian RBFs for 1-dimensional and 2-dimensional inputs.

A **radial basis function (RBF) network** is a type of artificial neural network, first introduced by Broomhead and Lowe (1988), that uses radial basis functions as activation functions. The network output is a combination of radial basis functions of the inputs and neuron parameters. An RBF network consists of three distinct layers: an input layer, a hidden layer that employs a non-linear radial basis function (RBF) as its activation function, and a linear output layer. The hidden layer performs a nonlinear mapping from the input space into a higher-dimensional space using a Gaussian or another kernel function. The output layer performs a weighted sum with a linear output. Figure 4.4 illustrates the architecture of an RBF network with an input layer, a hidden layer, and an output layer.

In an RBF network using Gaussian activation functions, the output of the j-th hidden unit for an n-dimensional input vector $\mathbf{x}$ is given by:

$$\phi_j(\mathbf{x}) = \exp\left(-\sum_{i=1}^{n} \frac{(x_i - c_{ij})^2}{b_{ij}^2}\right) \tag{4.1}$$

where $\mathbf{c}_j$ is the center vector of the j-th hidden unit with components c_{ij}, and $\mathbf{b}_j$ is the spread vector with components b_{ij}. The components of the spread vector $\mathbf{b}_j$ are related to the original spread parameters of the Gaussian function by $b_{ij} = \sqrt{2}\sigma_{ij}$.

The output of a multi-input, multi-output RBF network is given by:

$$y_k(\mathbf{x}) = \sum_{j=1}^{n_0} w_{jk}\phi_j(\mathbf{x}) \quad \forall k \in \{1, 2, \ldots, n'\} \tag{4.2}$$

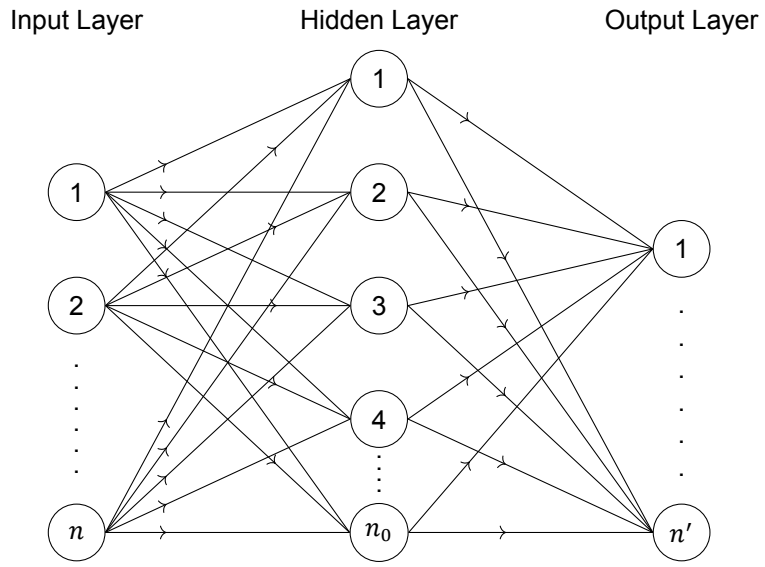


Figure 4.4: Architecture of a feed-forward Radial Basis Function (RBF) Network. This diagram illustrates the three-layer structure consisting of an input layer, a hidden layer with RBF activation functions, and a linear output layer. The network shown has n input nodes, n_0 hidden nodes, and n' output nodes, demonstrating its capability to map multiple inputs to multiple outputs. This architecture forms the basis of the policy functions used in our EMODPS approach for the Lower Omo-Gibe Basin (Broomhead and Lowe, 1988).

where $\mathbf{x}$ is an n-dimensional input vector, n_0 is the number of hidden units, $\phi_j(\mathbf{x})$ is the output of the j-th hidden unit, and w_{jk} is the weight associated with the j-th hidden unit for the k-th output.

Radial basis function networks are utilized in various applications, including image recognition, data interpolation, anomaly detection, and robotics control (Schwenker et al., 2001). Within the DPS framework, RBF networks are used as parameterized policy functions (Giuliani et al., 2020).

Depending on the problem's nature, parameter space complexity, and computational resources, the parameters of the RBF network can be optimized using gradient-based methods, evolutionary algorithms, or other optimization techniques. The non-differentiability and complex objective functions in water resource management problems make evolutionary algorithms a suitable choice for optimizing RBF networks within the EMODPS framework. Moreover, evolutionary algorithms excel in exploring large, complex, and multi-modal parameter spaces. They are more likely to escape local minima and find global optima in such spaces (Goldberg, 1989; Eiben and Smith, 2003; Mota et al., 2012).

Multi-Objective Evolutionary Algorithms (MOEAs) are stochastic search methods simulating natural selection and biological evolution. MOEAs use operators mimicking mating, mutation, and selection processes to solve multi-objective problems, starting from an initial population of randomly generated solutions and iteratively improving them (Coello et al., 2007).

NSGA-III (Nondominated Sorting Genetic Algorithm III) is an advanced multi-objective optimization algorithm designed to address the challenges of many-objective problems (i.e., problems with more than three objectives) (Deb and Jain, 2014; Jain and Deb, 2014). Similar to its predecessor, NSGA-II, NSGA-III employs a nondominated sorting approach to classify solutions based on Pareto dominance and maintains an elitist principle to ensure the best solutions are carried forward to the next generation (Deb et al., 2002). However, unlike NSGA-II, which relies on crowding distance to maintain diversity, NSGA-III uses a set of predefined reference points to guide the selection process, ensuring well-distributed solutions along the Pareto front. This makes NSGA-III particularly effective for higher-dimensional objective spaces. Figure 4.5 illustrates the key steps of NSGA-II and NSGA-III algorithms.

The key innovation of NSGA-III lies in its reference point-based selection mechanism, which enhances its ability to scale with an increasing number of objectives (Deb and Jain, 2014; Jain and Deb, 2014). During environmental selection, NSGA-III assigns solutions to reference points and selects those closest to each reference point, ensuring uniform distribution across the Pareto front. This approach improves performance in handling many-objective problems and addresses the computational inefficiencies of crowding distance calculations in high dimensions (Cheng et al., 2016). Despite these

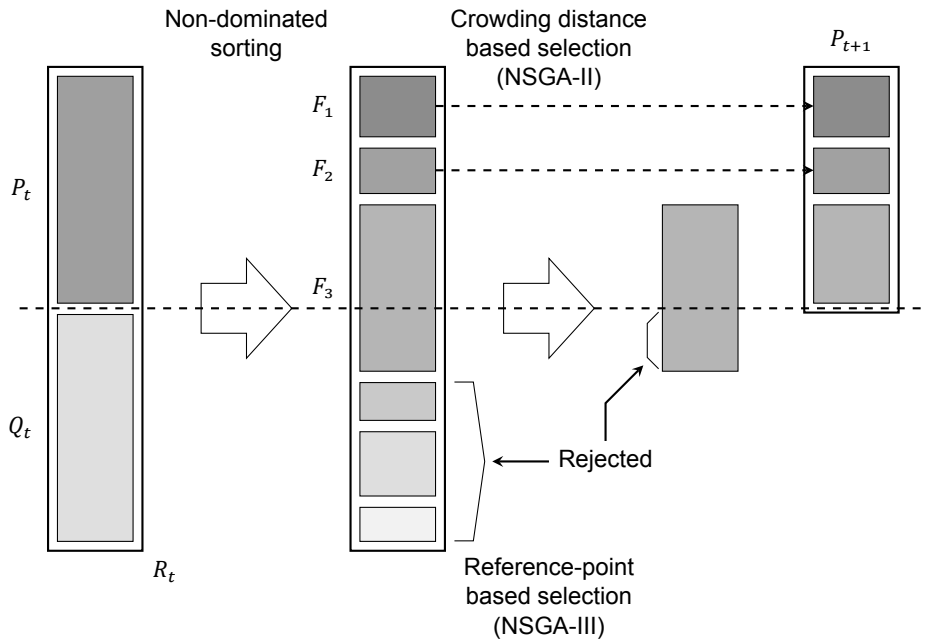


Figure 4.5: Schematic Diagram of NSGA-II and NSGA-III algorithms. This figure illustrates the key steps in these advanced multi-objective optimization algorithms, including non-dominated sorting, selection processes, and the creation of new populations. The diagram highlights the differences between NSGA-II's crowding distance-based selection and NSGA-III's reference point-based approach, demonstrating their respective strategies for maintaining diversity in high-dimensional objective spaces (Deb et al., 2002; Deb and Jain, 2014; Jain and Deb, 2014).

differences, NSGA-III retains the core mechanisms of NSGA-II, such as binary tournament selection and fast nondominated sorting, building on a solid foundation. Consequently, NSGA-III has been widely adopted in fields such as engineering design, bioinformatics, and environmental management, where optimizing multiple conflicting objectives is crucial.

In the present study, the Non-dominated Sorting Genetic Algorithm III (NSGA-III) is employed for evolutionary multi-objective optimization within the EMODPS framework since it is well-suited for three objectives.

We use the hypervolume indicator to assess the performance and convergence of MOEAs. This metric quantifies the volume of the objective space dominated by the solutions within the population, calculated with respect to a reference point. As the algorithm progresses and solutions approach the Pareto front, the hypervolume increases, making it a reliable measure of convergence and solution spread. Our methodology for convergence analysis involves the following:

- Computing the hypervolume for each generation.
- Plotting hypervolume against the number of function evaluations (NFEs).
- Repeating for multiple runs with different random seeds to assess consistency and robustness.

The resulting convergence plots provide insights into the algorithm's performance, including exploration effectiveness, solution quality improvement, and consistency across different initial conditions. Specific results and analysis will be presented in the Results chapter.

4.2.2. Applying EMODPS Framework in the Lower Omo-Gibe Basin

As discussed previously, in our case study, we aim to optimize the operation of the Gibe III and Koysha reservoirs and the Kuraz Headworks in the Lower Omo-Gibe Basin. The decision variables are the release decisions from the Gibe III and Koysha reservoirs and the percentage of inflow at Kuraz Headworks released to the irrigation canals. These decisions are made based on the storage levels in the reservoirs and the time index.

Problem Formulation
The release decisions are represented as follows:

- $R_{\text{Gibe-III}}(t)$ - Release from Gibe III dam
- $R_{\text{Koysha}}(t)$ - Release from Koysha dam
- $R_{\text{Irrigation}}(t)$ - Percentage of inflow at Kuraz Headworks released to irrigation canals

The release decisions must be computed at each time step t in the simulation horizon. In our case study, we consider a simulation horizon of 12 years with an integration time step of 12 hours.

Using the EMODPS framework, we base the release decisions on the storage levels in the reservoirs and the time index. The storage levels in the Gibe III and Koysha reservoirs are represented as follows:

- $V_{\text{Gibe-III}}(t)$ - Storage in Gibe III reservoir
- $V_{\text{Koysha}}(t)$ - Storage in Koysha reservoir
- $\tau(t)$ - Time Index

The storage levels are in m^3 and the time index (i.e.τ) for an integration time step is the month of the year ranging from 1 to 12 at the start of the integration time step.

We use the RBF network using Gaussian activation functions to parameterize the policy function that maps the inputs to the decisions. The inputs to the RBF network are normalized as follows:

- $x_1 = \frac{V_{\text{Gibe-III}}(t) - V_{\text{Gibe-III,min}}}{V_{\text{Gibe-III,max}} - V_{\text{Gibe-III,min}}}$
- $x_2 = \frac{V_{\text{Koysha}}(t) - V_{\text{Koysha,min}}}{V_{\text{Koysha,max}} - V_{\text{Koysha,min}}}$
- $x_3 = \frac{\tau(t) - \tau_{\text{min}}}{\tau_{\text{max}} - \tau_{\text{min}}}$

The normalization ensures that the inputs are scaled to the range $[-1, 1]$.

As we need to make three decisions, we have three outputs from the RBF network:

- y_1 - Normalized release from Gibe III dam
- y_2 - Normalized release from Koysha dam
- y_3 - Normalized percentage of inflow at Kuraz Headworks released to irrigation canals

RBF Network Architecture
As we have three inputs (i.e. x_1, x_2, x_3) and three outputs (i.e., y_1, y_2, y_3), the RBF network is designed to have six hidden units. This is inspired by the literature on RBF networks for water resource systems (Giuliani et al., 2016), which recommends that for n inputs and n' outputs, the RBF network should have $n + n'$ hidden units. As a result, the RBF network has $3 + 3 = 6$ RBF units in the hidden layer. The relation between inputs and outputs is represented as follows:

$$
\begin{aligned}
y_1(\mathbf{x}) &= \sum_{j=1}^{6} w_{j1} \exp\left(- \sum_{i=1}^{3} \frac{(x_i - c_{ij})^2}{b_{ij}^2} \right) \\
y_2(\mathbf{x}) &= \sum_{j=1}^{6} w_{j2} \exp\left(- \sum_{i=1}^{3} \frac{(x_i - c_{ij})^2}{b_{ij}^2} \right) \\
y_3(\mathbf{x}) &= \sum_{j=1}^{6} w_{j3} \exp\left(- \sum_{i=1}^{3} \frac{(x_i - c_{ij})^2}{b_{ij}^2} \right)
\end{aligned}
\tag{4.3}
$$

where $\mathbf{x} = (x_1, x_2, x_3)$ is the input vector, $y_k(\mathbf{x})$ is the output for the k-th output, w_{jk} are the weights, c_{ij} are the centers, and b_{ij} are the spreads of the RBF units. The centers take values in the range

$(-1, 1]$, and the spreads are positive real numbers in the range $[0, 1]$. The weights are non-negative real numbers in the range $[0, 1]$ such that the weights for each output sum to 1, i.e., $\sum_{j=1}^{6} w_{jk} = 1$ for $k \in \{1, 2, 3\}$.

The above setup ensures that the outputs of the RBF network are always in the range $[0, 1]$. These outputs are denormalized to determine the release decisions at each time step. The inverse normalization is given by:

- $R_{\text{Gibe-III}}(t) = y_1 \times (R_{\text{Gibe-III,max}} - R_{\text{Gibe-III,min}}) + R_{\text{Gibe-III,min}}$
- $R_{\text{Koysha}}(t) = y_2 \times (R_{\text{Koysha,max}} - R_{\text{Koysha,min}}) + R_{\text{Koysha,min}}$
- $R_{\text{Irrigation}}(t) = y_3 \times (R_{\text{Irrigation,max}} - R_{\text{Irrigation,min}}) + R_{\text{Irrigation,min}}$

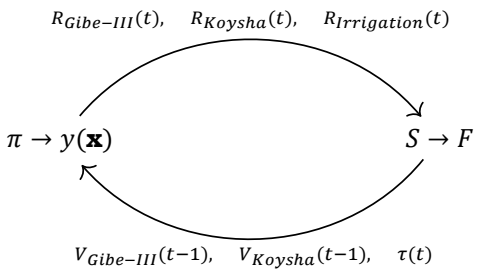
Since the third output is a percentage, the value for $R_{\text{Irrigation, max}}$ is 100, and $R_{\text{Irrigation, min}}$ is 0. The values for $R_{\text{Gibe-III, min}}$ and $R_{\text{Koysha, min}}$ are set to 0, and the maximum values are determined based on the maximum turbine flow rates at the respective dams. As a result the units for the first two outputs are in m^3/s and the third output is a percentage. As discussed later, the release decisions resulting from policy functions are subject to physical and operational constraints. The actual releases are computed by accounting for mass balance equations and maximum and minimum reservoir levels (if applicable). As a result, the actual releases may be higher or lower than the release decisions from the RBF network. The exact mechanism for computing the actual releases is discussed in the next section.

The release decision for the irrigation canals at Kuraz Headworks is determined by taking a product of inflow at the headworks and $R_{\text{Kuraz}}(t)$. In our model the inflow at Kuraz Headworks is basically the actual release from the Koysha reservoir. While the release to canals (in terms of m^3/s) could also be taken as an output of the RBF network, it was found that the RBF network was not able to capture the dynamics of the Headworks effectively since the inflow to the headworks is directly dependent on the release from the Koysha reservoir. Moreover, the release to canals directly affects the remaining flow that is released to the main channel. As a result, the release to the irrigation canals is determined using a simple linear relationship with the release from the Koysha reservoir, where the release decision $R_{Irrigation}$ (in terms of percentage) becomes an independent decision variable that can freely vary between 0 and 100. For instance, if the release from Koysha reservoir is 100 m^3/s, and the release to canals is 30%, then the release to the canals is 30 m^3/s. The remaining 70% (i.e.70 m^3/s) is released to the main channel.

In our RBF network, we have a total of $6{\times}3 = 18$ centers (i.e. number of hidden units×number of inputs), $6 \times 3 = 18$ spreads (i.e. number of hidden units × number of inputs), and $6 \times 3 = 18$ weights (i.e. number of hidden units × number of outputs), for a total of 54 parameters to be optimized.

Policy Parameters and Objective Functions

Based on the above setup, we can understand a policy π as a vector of 54 parameters that determine the release decisions from the Gibe III and Koysha reservoirs and the Kuraz Headworks at each time step in a simulation of the Lower Omo-Gibe Basin. The release decision determines the time series of state variables (i.e., reservoir storage levels and flow rates). These state variables determine the hydropower generation, irrigation demand deficit, environmental demand deficit, etc. For instance, the water lost to evaporation from the reservoirs will be determined by the time series of storage levels in the reservoirs and the seasonality of the evaporation rates. The time series of releases from the reservoirs and the storage levels in the reservoirs will determine the hydropower generation. This mechanism can we represented as follows:



where π is the policy (i.e., 54 parameters of the RBF network), $y(\mathbf{x})$ is the RBF network with parameters π, S is the time series of state variables, and F is the vector of objective scores (3-dimensional in our case). The loop in the above diagram is the same as the *Simulation Loop* shown in Figure 4.2.

The objectives to be optimized are as follows:

- $F_1(S)$ - Mean Power Output from Gibe-III and Koysha Plants
- $F_2(S)$ - 90$^{\text{th}}$ Percentile of Monthly Irrigation Demand Deficits
- $F_3(S)$ - 90$^{\text{th}}$ Percentile of Monthly Environmental Flow Deficits

The objectives are evaluated based on the time series of state variables (i.e., storage levels in the reservoirs, flow rates and demand rates). The objectives are computed as follows:

- $F_1(S) = \frac{1}{T}\sum_{t=1}^{T} P_{\text{Gibe-III}}(t) + P_{\text{Koysha}}(t)$
- $F_2(S) =$ 90th percentile of $\left(\max\left(0, 100 \times \frac{D_{\text{Irrigation}}(m) - \overline{Q_{\text{Irrigation}}}(m)}{D_{\text{Irrigation}}(m)}\right)\right)$
- $F_3(S) =$ 90th percentile of $\left(\max\left(0, 100 \times \frac{D_{\text{Environment}}(m) - \overline{Q_{\text{Environment}}}(m)}{D_{\text{Environment}}(m)}\right)\right)$

where T is the total number of time steps in the simulation horizon, $P_{\text{Gibe-III}}(t)$ and $P_{\text{Koysha}}(t)$ are the power outputs from Gibe-III and Koysha plants, $\overline{Q_{\text{Irrigation}}}(m)$ and $\overline{Q_{\text{Environment}}}(m)$ are the average releases for irrigation and environment in month m, and $D_{\text{Irrigation}}(m)$ and $D_{\text{Environment}}(m)$ are the demand rates for irrigation and environment in month m. The deficit rates are then calculated for each month in the simulation horizon and sorted to find the 90th percentile.

The power generation from the Gibe III and Koysha plants is determined by the flow rates released from the reservoirs and the head at the turbines. The power generation is computed using the following equations:

- $P_{\text{Gibe-III}}(t) = \eta_{\text{Gibe-III}} \times \rho \times g \times \left(H_{\text{Gibe-III}}(t) - H_{\text{turbine, Gibe-III}}\right) \times Q_{\text{Gibe-III}}(t)$
- $P_{\text{Koysha}}(t) = \eta_{\text{Koysha}} \times \rho \times g \times \left(H_{\text{Koysha}}(t) - H_{\text{turbine, Koysha}}\right) \times Q_{\text{Koysha}}(t)$

where $\eta_{\text{Gibe-III}}$ and η_{Koysha} are the efficiencies of the Gibe III and Koysha plants, ρ is the density of water, g is the acceleration due to gravity, $H_{\text{Gibe-III}}(t)$ and $H_{\text{Koysha}}(t)$ are the heights of the water at Gibe III and Koysha reservoirs at time t, $H_{\text{turbine, Gibe-III}}$ and $H_{\text{turbine, Koysha}}$ are the heights of the turbines at Gibe III and Koysha plants, and $Q_{\text{Gibe-III}}(t)$ and $Q_{\text{Koysha}}(t)$ are the flow rates released from Gibe III and Koysha reservoirs. $H_{\text{Gibe-III}}(t)$ and $H_{\text{Koysha}}(t)$ are determined using bathymetry and the storage levels in the reservoirs (i.e. $V_{\text{Gibe-III}}(t)$ and $V_{\text{Koysha}}(t)$).

The values for $Q_i t$ and $V_i t$ for all reservoirs and flows are available in the time series of state variables S. Therefore, the vector F with objective scores (i.e. F_1, F_2, and F_3) can be seen as functions of the time series of state variables S.

Searching Pareto-Approximate Solutions

As discussed earlier, we use NSGA-III algorithm to search for Pareto-approximate solutions. The optimization procedure, depicted as *Optimization loop using Direct Policy Search* in Figure 4.2 generates the Pareto-approximate solutions. The mechanism for optimization can be summarized as follows:

1. Initialize a population of Policy functions (i.e., different sets of 54 parameters, each defining an RBF network).
2. Simulate the system for each policy function in the population to obtain the time series of state variables.
3. Evaluate the objectives for each policy function based on the time series of state variables (using the simulation loop discussed earlier).
4. Keep the Policy functions that perform well based on the objective scores and discard the rest.
5. Generate new Policy functions by mutating and crossing over those performing well.

6. Combine the policy functions from the previous two steps to form the new population.
7. Repeat steps 2 to 6 for a fixed number of generations or until convergence.
8. The final set of Policy functions, with dominated ones removed, represent the Pareto-approximate solutions.

The Pareto-approximate solutions represent the trade-offs between the objectives. The results for the optimization of reservoir operations in the Lower Omo-Gibe Basin using the EMODPS framework will be discussed in next chapter. The rest of the chapter will focus on the design, implementation, and application of the HydroWizard framework for simulating and optimizing water resource systems.

4.3. HydroWizard: An Advanced Framework for Water Resource Systems Modeling

HydroWizard is a comprehensive framework for efficient simulation, optimization, and analysis of water resource systems. It addresses the reproducibility crisis in environmental modeling (Choi et al., 2023) by providing a reproducible, transparent, scalable, and extensible environment. HydroWizard extends the FAIR principles (Wilkinson et al., 2016) to the modeling process itself, integrating system specification, execution, visualization, and data management components.

The HydroWizard framework is conceptualized as a language-agnostic paradigm, designed with the flexibility to be implemented across diverse programming environments. To empirically validate its practicality and demonstrate its efficacy in real-world scenarios, we have developed a full-fledged implementation in Python. This choice of language allows for rapid prototyping and wide accessibility within the scientific community. The open-source nature of this implementation aligns with principles of scientific transparency and reproducibility, and is accessible through the following channels:

- Source code: `https://github.com/yugdeep/hydrowizard`
- Package: `https://pypi.org/project/hydrowizard/`
- Documentation: `https://hydrowizard.readthedocs.io/`

In this study, we employ the framework to model the Lower Omo-Gibe Basin and analyze trade-offs in Pareto-optimal solutions for reservoir operations, demonstrating its capability to handle complex water resource systems.

This section explores the framework's design principles, model specification language, execution engine, optimization module, data logging mechanism, visualization tools, and user interface options. We conclude with an application to the Lower Omo-Gibe Basin, aiming to demonstrate HydroWizard's potential to advance water resource modeling and contribute to more informed, efficient, and sustainable water management practices.

4.3.1. Design Principles

HydroWizard's development is guided by the following key principles:

- **Enhancing Scientific Reproducibility:** By providing a transparent and consistent environment for water resource modeling, HydroWizard aims to facilitate the sharing and verification of models, data, and results within the scientific community.
- **Promoting Methodological Transparency:** The framework encourages clear articulation of assumptions, data sources, and methodologies, fostering comprehensive scrutiny and understanding of models by researchers and stakeholders.
- **Enabling System Scalability:** HydroWizard is designed to accommodate large-scale water resource systems, supporting complex models with multiple components, objectives, and constraints across diverse modeling scenarios.
- **Facilitating Framework Extensibility:** The modular architecture allows for the incorporation of custom components, algorithms, and visualizations, ensuring adaptability to evolving research needs and technological advancements.

- **Optimizing User Accessibility:** By offering multiple user interfaces, including a command-line interface (CLI) and a Python package, HydroWizard caters to users with varying levels of expertise.
- **Maximizing Computational Efficiency:** Through the automation of mass balancing and physical constraint satisfaction, coupled with robust optimization and simulation capabilities, HydroWizard streamlines the model setup process and reduces the need for extensive programming skills.

These design principles underpin HydroWizard's comprehensive approach to water resource modeling, enabling researchers and practitioners to efficiently explore diverse model configurations, identify optimal solutions, simulate system behavior, and visualize results. This framework facilitates a more thorough exploration of the modeling space, potentially yielding deeper insights into policy implications for water resource management.

4.3.2. Model Specification Language

The HydroWizard framework introduces a robust and flexible model specification language designed to facilitate the definition of complex water resource systems. This language serves as a bridge between human-interpretable system descriptions and machine-executable models, enabling efficient computation while maintaining accessibility for users with varying levels of technical expertise. For a comprehensive reference guide to the language, readers are directed to Appendix A.

Conceptual Framework

At its core, the HydroWizard Model Specification Language represents water resource systems as Directed Acyclic Graphs (DAGs) composed of nodes and flows. This structure provides a versatile foundation for modeling diverse system components and their interactions. The primary elements of the language are:

- **Nodes**: Represent discrete points in the system, such as reservoirs, confluences, and measurement stations.
- **Flows**: Represent water transfers between nodes, including inflows, outflows, releases, and evaporation.

Nodes are categorized as either reservoir nodes (with storage capacity) or non-reservoir nodes. Flows are classified into three types based on their determination method:

- **X Flows**: External, typically stochastic inflows.
- **L Flows**: Decision-dependent flows, representing system controls.
- **R Flows**: System-dependent flows, governed by mass balance equations.

The specification language also incorporates basin-wide parameters and optimization objectives, providing a comprehensive framework for system modeling and analysis.

Implementation Architecture

HydroWizard employs YAML (YAML Ain't Markup Language) as the underlying format for model specifications. This choice leverages YAML's readability, simplicity, and support for complex data structures. The YAML-based specification allows for:

- Clear and intuitive definition of system components and relationships.
- Hierarchical representation of model elements.
- Inclusion of comments for enhanced documentation and interpretability.

For detailed syntax, field descriptions, and examples of the YAML-based model specification, please refer to Appendix A.

Key Features and Capabilities
The HydroWizard Model Specification Language offers several notable features:

- **Flexibility**: Supports both constant and time-varying parameter values.
- **Scalability**: Capable of representing systems ranging from simple, small-scale models to complex, large-scale river basins.
- **Extensibility**: Allows for the definition of custom objective functions to meet specific modeling requirements.
- **Interoperability**: The machine-readable format facilitates integration with various simulation and optimization tools.

An exemplar model specification, demonstrating the structure and key components of a HydroWizard configuration, is provided in Listing 1. For a more comprehensive example applied to a real-world water resource system, readers are encouraged to examine the complete configuration file for the Lower Omo-Gibe Basin case study in Appendix D.

The HydroWizard Model Specification Language thus provides a powerful and accessible tool for water resource system modeling, enabling researchers and practitioners to efficiently define complex hydrological systems.

4.3.3. Model Execution Engine

The model execution engine exploits the power of model specification language to automate everything from model setup to simulation and optimization. It is designed to efficiently parse the model specifications, search for pareto optimal solutions using the specified optimization methods, and simulate the system behavior over time.

In the HydroWizard framework, the modeler's job is complete once a water resource system's specifications have been defined in the model specification language (in a YAML file). The model execution engine takes care of the rest, ensuring that the model is correctly parsed, the optimization methods are effectively applied, and the simulation results are accurately captured. The Python package for HydroWizard already provides a rich set of functionalities for model execution, including simulation, optimization, and visualization tools. There is also support for parallel computing, which enables the framework to handle large-scale simulations and optimizations efficiently. Limited debugging support is also provided to help users identify and resolve issues in the model specification. Additional features can be easily added to the package to enhance its capabilities and extend its applicability to different types of water resource systems.

HydroWizard framework allows practitioners to model a water resource system in a high-level language, while the model execution engine takes care of the technical details of setting up and running the optimization and simulations efficiently, accurately, and reproducibly. This separation of concerns ensures that the model development process is streamlined and the results are reliable and consistent. It allows researchers and practitioners to focus on analyzing and interpreting the results rather than getting bogged down in the technicalities of model setup and execution. At the same time, it allows technical experts to develop and extend the framework collaboratively, adding new features and functionalities to meet the evolving needs of the wider water resource modeling community.

In the following sections, we will discuss the key components of the model execution engine and how they work together to simulate and optimize water resource systems.

System Architecture
The model execution engine in HydroWizard leverages a Directed Acyclic Graph (DAG) to represent the water resource system. DAG complements the model specification language by providing a structured and efficient way to organize the system components and interactions.

A DAG is a graph structure with directed edges and no cycles, ensuring a unidirectional flow from sources to sinks (Cormen et al., 2009). This structure is highly suitable for modeling water resource systems, as it naturally represents the hierarchical and sequential nature of water flow and resource distribution.

The use of a DAG in HydroWizard provides several advantages:

- Clarity and Structure: DAGs offer a clear and intuitive way to represent the components and interactions within a water resource system. Nodes represent system components such as reservoirs, rivers, and demand sites, while edges (flows) represent the interactions and dependencies between these components.

- Efficiency: DAGs allow for efficient computation of simulations and optimizations as the acyclic nature prevents infinite loops and ensures that flow rate computations are processed in a logical sequence.

- Modularity and Extensibility: By organizing the model into nodes and flows, users can easily add, remove, or modify components without disrupting the overall system structure. This modularity aligns well with the extensible nature of HydroWizard.

DAG is a powerful abstraction that has been successfully applied in various domains, including machine learning, parallel computing, and workflow management. Some examples of DAG-based frameworks include:

- TensorFlow and PyTorch: These machine learning frameworks use DAGs to define and manage the flow of data through neural networks. Each operation in the network is a node, and the directed edges represent data dependencies (Abadi et al., 2016; Paszke et al., 2019).

- Apache Airflow: This platform uses DAGs to manage and orchestrate workflows, ensuring tasks are executed in a specified order without cycles, similar to how HydroWizard handles water flow and resource distribution (Foundation, 2015).

- Dask: This parallel computing library uses DAGs to optimize task scheduling and execution, enabling efficient processing of large datasets and complex computations (Rocklin, 2015).

By adopting a DAG-based architecture, HydroWizard ensures that the water resource modeling process is both efficient and scalable, leveraging the same principles that underpin some of the most advanced computational frameworks in use today.

Figure 4.6 shows an example of a water resource system (Lower Omo Basin in this case, as specified in Appendix A) represented as a DAG. The rectangular nodes represent reservoir nodes with storage capacities, the circular nodes represent non-reservoir nodes such as confluences and head regulators. The green dots represent dummy nodes that are used as source nodes for external flows entering the system. The red dots represent dummy nodes that are used as target nodes for flows leaving the system. The directed edges represent the flows between the nodes, defining the flow paths in the system. The green edges represent X flows (external inflows), the blue edges represent L flows (decision-dependent flows), and the red edges represent R flows (system-dependent flows). This visual representation provides an intuitive way to understand the structure and interactions within the water resource system and will be expanded further in the subsequent sections to include additional components and details.

Parsing Mechanism

The HydroWizard parser mechanism translates the model specification language into a Directed Acyclic Graph (DAG) representing the water resource system. This involves parsing the YAML file, extracting information, and constructing nodes, flows, and basin objects. The parser validates the input, ensuring correct formatting and handling errors. If the basin is not connected, it raises an error for the user to fix. The parser also generates a visual representation of the system structure for verification in form of an annotated DAG graph as shown in Figure 4.6.

Utilizing the PyYAML library, the parser extracts components from the YAML file and constructs the DAG via the Basin object, which then creates Node and Flow objects. The Basin object represents the top-level system and provides methods for simulation and optimization.

Designed for parallelization, the parser facilitates the efficient processing of large-scale models. In the EMODPS approach, optimization often runs in parallel across multiple cores or nodes. The parser quickly constructs the Basin object by taking the path of the model specification file as input and returning the Basin object, avoiding the slow process of pickling large objects for parallel execution.

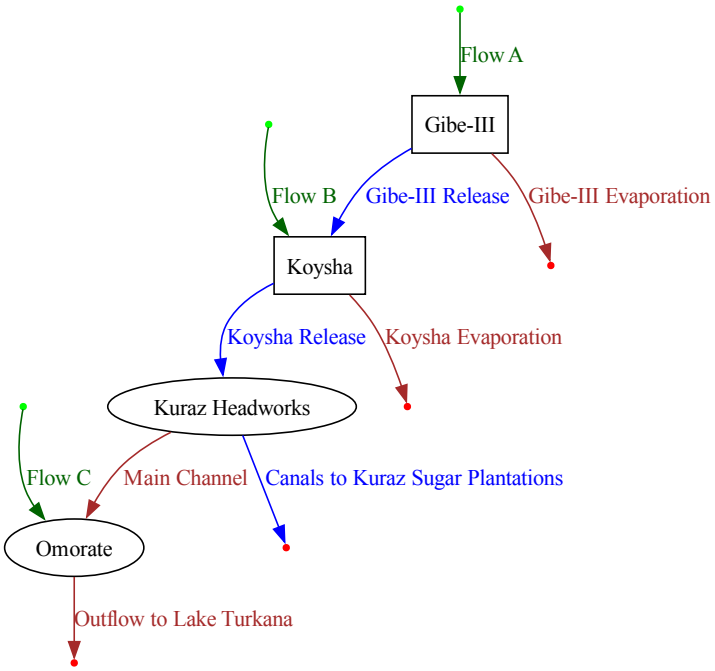


Figure 4.6: Directed Acyclic Graph (DAG) representation of the Lower Omo-Gibe Basin water resource system. This graph illustrates the complex network of reservoirs, flows, and control points in the basin. Rectangular nodes represent reservoirs, circular nodes depict non-reservoir elements, and colored edges indicate different types of flows (green for external inflows, blue for decision-dependent flows, and red for system-dependent flows).

Simulation Process

The simulation process forms the core of the HydroWizard framework, dynamically executing the model to simulate the behavior of the water resource system over time. This process comprises several critical steps, each contributing to the accurate representation of system dynamics:

- **Initialization**: The simulation commences with the initialization of the system state, including reservoir storage levels, start time, and other pertinent variables. Initial conditions are established based on the model specification and user-defined parameters.
- **Time Stepping**: The simulation progresses through discrete time steps, with the system state updated at each interval based on inflows, outflows, and other system dynamics. The time step size is determined by the integration interval specified in the model configuration.
- **Release Decisions**: At each time step, release decisions from reservoirs and other control points are determined by policy functions, which take relevant state variables as input.
- **Flow Computation**: Flow rates within the system are computed based on release decisions, physical constraints of nodes and flows, and governing physical laws. This step enforces mass balance across the system, ensuring accurate accounting of all inflows, outflows, and storage changes at each time step.
- **Objective Function Calculation**: Upon completion of the simulation, the objective function, as defined in the model specification, is calculated based on the time series of state variables.

The initialization phase creates an empty time series of state variables S that encapsulates the system state at each time step. This includes storage levels in reservoirs and flow rates for each system flow. The state variables are categorized into stocks (reservoir volumes) and flows (flow rates). A matrix of dimensions $n_r \times (m+1)$ is generated for stocks, where n_r represents the number of reservoir nodes and m the number of time steps. Concurrently, a matrix of dimensions $n_f \times m$ is created for flows, with n_f denoting the number of flows in the system.

The time stepping process iterates over each time step, updating the system state and computing flow rates based on release decisions. These decisions are determined by policy functions, which can be rule-based, optimization-based, or machine learning-based. The current implementation of HydroWizard supports Radial Basis Function (RBF) network-based policy functions, optimized using the Evolutionary Multi-Objective Direct Policy Search (EMODPS) framework to identify Pareto-optimal solutions.

The flow computation step, central to the simulation process, determines flow rates at each time step for any arbitrary water resource system. This critical process ensures the accuracy and reliability of simulation results. The HydroWizard framework employs robust algorithms for mass balance maintenance and flow computation, guaranteeing accurate capture and simulation of system dynamics over time.

The flow computation process occurs in the following sequence:

1. **Assignment of X Flow Values**: External flows are assigned values based on time series data provided in the model specification.
2. **Pre-decision Assignment of R Flow Values**: System-dependent flows, such as evaporation and seepage, are computed based on relevant system variables and physical laws, independent of release decisions.
3. **Computation of Release Decisions**: Release decisions from reservoirs and other control points are determined using policy functions.
4. **Assignment of Actual L Flow and Remaining R Flow Values**: Decision-dependent flows are assigned values based on release decisions and system physical constraints. This step is further divided into three sub-steps:
   (a) Computation of min/max constraints on L-flows arising from dependent flow constraints using a backward pass in the Directed Acyclic Graph (DAG).

(b) Calculation of min/max constraints on L-flows derived from reservoir volume constraints.

(c) Assignment of missing flow rates using a forward pass in the DAG.

5. **Reservoir Volume Update**: Storage levels in reservoirs are updated based on computed inflow and outflow rates.

The intricacies of satisfying min/max constraints on flow rates and storage levels, as well as the assignment of actual flow rates to L-flows and remaining R-flows, are detailed in Algorithms 1 and 2 in Appendix B. These algorithms elucidate the sophisticated processes employed to ensure physical consistency and policy adherence in the system simulation.

Figure 4.7 provides a visual representation of the flow computation process in the HydroWizard framework. It effectively communicates the complexity of the flow computation process, showcasing the topological sorting of the system's DAG, the propagation of constraints, and the sequential assignment of flow rates. The figure is strategically divided into two main sections: the upper section (Algorithm 1) depicts the backward pass for propagating min/max constraints, while the lower section (Algorithm 2) illustrates the forward pass for assigning actual flow rates.

By employing this sophisticated flow computation process, the HydroWizard framework ensures that all physical and operational constraints are respected in any arbitrary water resource system. This approach enables the accurate simulation of complex water resource systems, providing a robust foundation for subsequent analysis and optimization of water management strategies.

4.3.4. Optimization Module

The optimization module in HydroWizard is tasked with identifying Pareto-optimal solutions that maximize or minimize the objectives defined in the model specification. The current Python implementation of HydroWizard incorporates the Evolutionary Multi-Objective Direct Policy Search (EMODPS) framework, which has been elaborated upon in preceding sections. This module offers ready-to-use implementations of Radial Basis Function (RBF) network-based policy functions, coupled with optimization algorithms for Pareto-optimal solution discovery. The number of basis functions is judiciously selected based on the dimensionality of decision variables and objectives.

The optimization process leverages the Pymoo library (Blank and Deb, 2020), which provides a comprehensive suite of multi-objective optimization algorithms, including NSGA-II, NSGA-III, MOEA/D, and RVEA. HydroWizard's Application Programming Interfaces (APIs) enable users to specify the optimization algorithm and configure parameters such as population size and number of generations. Furthermore, the optimization process supports parallel computing, allowing users to efficiently utilize computational resources and accelerate the search for Pareto-optimal solutions. This parallelization can be invoked by specifying the number of processes in the optimization call, as detailed in Appendix C.

4.3.5. Data Logging Mechanism

Given the potential for protracted convergence times in EMODPS optimizations, particularly for large-scale water resource systems, the HydroWizard incorporates a sophisticated data logging mechanism. This feature enables real-time tracking of the optimization process and facilitates on-the-fly analysis of results. The Python implementation exports the parameters of policy functions and their corresponding objective values at each generation to a CSV file.

Additionally, the Python package supports data export to a remotely hosted PostgreSQL database, allowing users to monitor the optimization process in real-time without direct server access. The remote logging module also facilitates the retrieval of policy functions and objective values from previous optimization runs. This functionality is particularly valuable for users seeking to resume an optimization run from a previous generation or to analyze results without manual data retrieval from the database.

HydroWizard's data logging mechanism ensures transparency, traceability, and reproducibility of the optimization process. The remote logging feature enhances flexibility and convenience, enabling multiple concurrent users to access optimization data ubiquitously. This capability is especially beneficial for collaborative research projects involving multiple researchers working on the same optimization problem or for remote monitoring of the optimization process.

Any policy function, represented as a list of parameter values for the RBF network, whether stored locally or fetched from the remote database, can be utilized in conjunction with the model specification

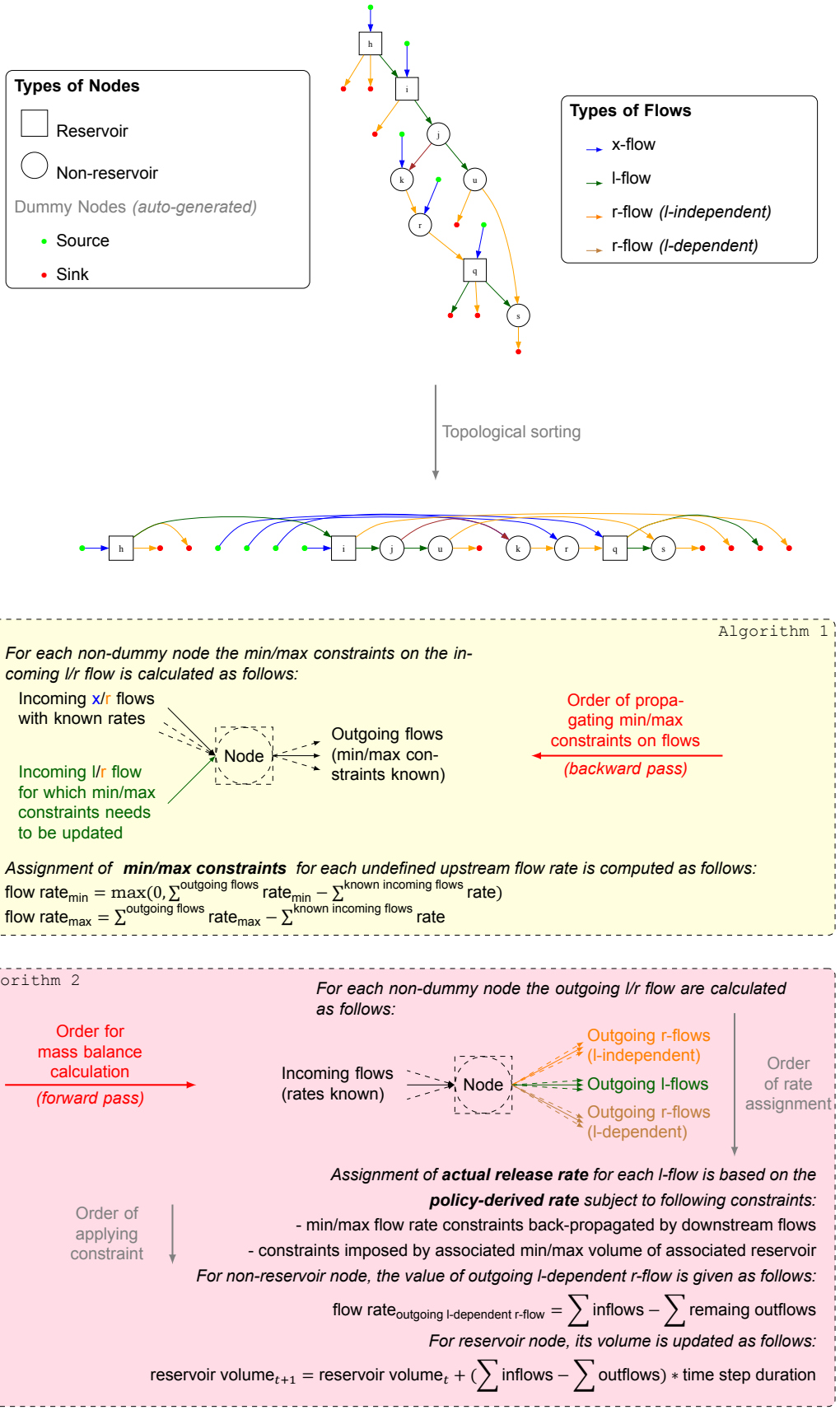


Figure 4.7: Flow computation process in the HydroWizard Framework. This comprehensive diagram illustrates the step-by-step procedure for computing flow rates in the water resource system, incorporating constraints, mass balance, and policy decisions. The figure is divided into two main sections: Algorithm 1 (in yellow) shows the backward pass for propagating min/max constraints, while Algorithm 2 (in pink) depicts the forward pass for assigning actual flow rates. This process ensures physical consistency and policy adherence in the system simulation.

file to simulate the water resource system. The HydroWizard Python package provides a dedicated function for this purpose. Simulation results, including time series of state variables, power generation, and demand deficits, can be optionally exported to CSV files for further analysis and visualization.

The data logging mechanisms, particularly the remote logging feature, facilitate the sharing of optimization results through web-based dashboards. This capability is instrumental in presenting results to stakeholders, decision-makers, and the wider research community in an accessible and interactive manner.

4.3.6. Visualization Tools

HydroWizard's Python implementation incorporates a comprehensive visualization tool suite that enables users to conduct graphical analyses of simulation results. These tools provide invaluable insights into system behavior, performance metrics, and the impacts of various management policies. A notable innovation in this suite is the development of state graphs, which extend the Directed Acyclic Graph (DAG) representation of the water resource system (as illustrated in Figure 4.6) to include state variables at each node and flow in the system at a given time step.

Figure 4.8 presents an exemplar state graph for the Lower Omo-Gibe Basin at a specific simulation time step. This visualization provides a comprehensive view of system dynamics and interactions, enabling users to discern system behavior and identify potential anomalies or unexpected outcomes.

The state graph incorporates reservoir storage levels, flow rates, and additional metrics such as power generation and reservoir head. The color-coding scheme distinguishes between external inflows (green), inputs for the RBF network (orange), release decisions derived from the RBF network outputs (blue), and system-derived flows and state variables (red).

These state graphs enhance transparency and accountability in the modeling process, allowing modelers and stakeholders to visualize the system state at any time step and identify discrepancies or inconsistencies in simulation results. The visualization module in HydroWizard's Python package provides functionality to generate state graphs for individual time steps or a series of time steps in the simulation process. This capability is particularly valuable for analyzing the impact of different management policies, evaluating system performance, and identifying areas for improvement in the water resource system.

4.3.7. User Interface Options

Installation of HydroWizard's Python package offers two primary interface options to cater to users with varying levels of expertise and requirements:

Command-Line Interface (CLI)
The Command-Line Interface (CLI) facilitates the execution of experiments through direct command inputs. It leverages the simulation and optimization scripts of the HydroWizard Python package, enabling efficient experimentation without necessitating in-depth knowledge of the underlying code structure. Detailed CLI commands for optimization and simulation processes are provided in Appendix C. These commands exemplify the framework's user-friendly approach to complex modeling tasks, allowing researchers to conduct sophisticated analyses without extensive engagement with the underlying code structure.

Python API
The HydroWizard Python package provides advanced users with direct access to the frameworks's functionalities. This interface is ideal for users with Python programming experience who seek to extend the framework's capabilities or integrate it with other tools and workflows. The package offers a high degree of flexibility and customization, enabling users to tailor the framework to their specific research needs.

4.4. Application to the Lower Omo-Gibe River Basin

This section delineates the application of the HydroWizard framework to the Lower Omo-Gibe River Basin, demonstrating its efficacy in modeling and optimizing complex river basin systems. The framework's application encompasses two primary components: model configuration and computational setup, each crucial for ensuring accurate representation and efficient optimization of the basin's dynamics.

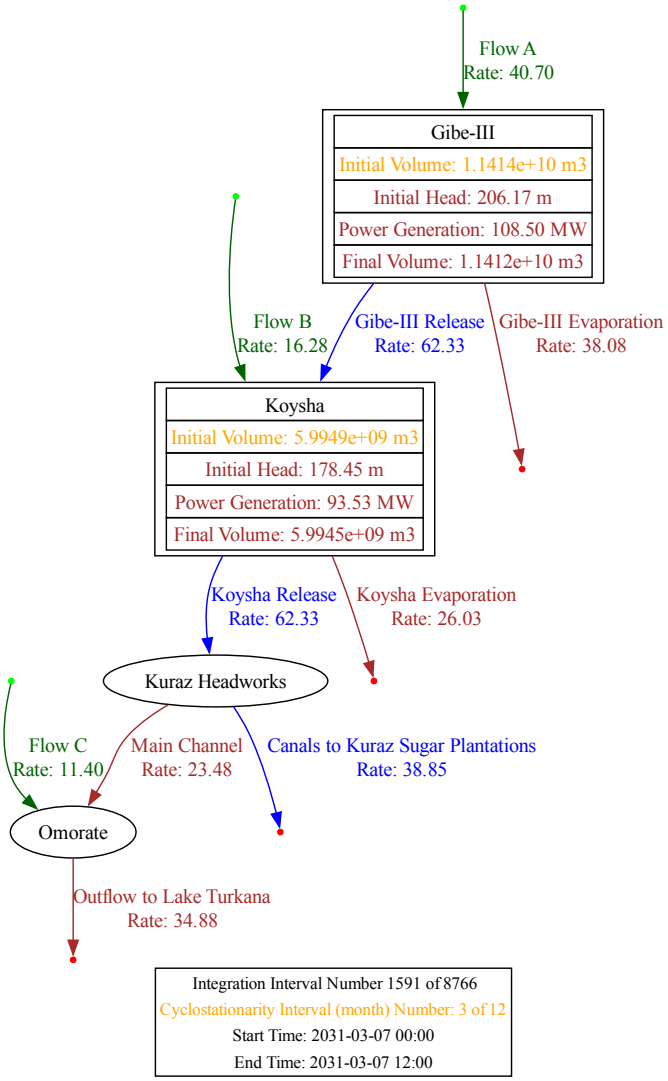


Figure 4.8: State graph representation of the Lower Omo-Gibe Basin water resource system at a specific simulation time step. This visualization extends the basic Directed Acyclic Graph (DAG) structure, including dynamic state variables and flow rates. Color coding differentiates between external inflows (green), policy-derived releases (blue), and system-computed variables (red). Orange variables represent inputs to the Radial Basis Function network. Reservoir levels, flow rates, power generation, and other key metrics are displayed, providing a comprehensive snapshot of the system's state. This state graph approach enables detailed analysis of system behavior, policy impacts, and potential anomalies at any given point in the simulation, enhancing transparency and facilitating in-depth evaluation of water management strategies.

4.4.1. Model Configuration

The configuration of the HydroWizard model for the Lower Omo-Gibe River Basin was accomplished through a structured YAML file, which serves as a comprehensive representation of the basin's characteristics. This configuration method facilitates a systematic approach to model definition, encompassing critical elements such as reservoir properties, hydropower plant specifications, environmental flow requirements, irrigation demands, and hydrological inputs.

The complete configuration file, provided in Appendix D, is a definitive reference for the model's structure and parameters.

4.4.2. Computational Setup

The computational implementation of the HydroWizard framework for the Lower Omo-Gibe River Basin optimization problem was designed to balance computational efficiency with solution quality. The optimization process was parameterized as follows:

- Simulation Horizon: 12 years
- Integration Time Step: 12 hours
- Number of Function Evaluations: 80,000

These parameters were selected based on preliminary analyses to ensure adequate temporal resolution and sufficient exploration of the solution space while maintaining computational feasibility.

A series of experiments were conducted to determine the most efficient configuration to optimize the performance of the multi-objective evolutionary algorithm employed by the framework. These experiments varied key algorithm parameters, including population size, number of generations, and degree of parallelization. The results of these experiments, detailed in Appendix E, led to the adoption of the following configuration:

- Population Size: 128
- Number of Generations: 625
- Number of Parallel Processes: 64

This configuration achieved a CPU utilization of 96.15%, demonstrating highly efficient use of the computational resources available on the DelftBlue node (Delft High Performance Computing Centre , DHPC) employed in this study.

4.4.3. Results Replication and Data Availability

In adherence to the principles of scientific reproducibility and transparency, we present a generalized methodology for replicating results derived from any water resource system modeled using the HydroWizard framework. This approach ensures that our findings, as well as those of future studies utilizing this framework, can be independently verified and further explored by the scientific community.

The cornerstone of this replication process is the sharing of a human-readable model configuration file, which encapsulates the entire structure and parameters of the water resource system under study. For our specific investigation of the Omo-Gibe basin, we have made the raw data from our optimization and simulation processes, along with the model configuration, publicly accessible at `https://github.com/yugdeep/omo-gibe`. To account for stochastic variability inherent in the optimization process, we conducted 10 repetitions using distinct random seeds.

The replication procedure, applicable to our study and extendable to any HydroWizard-based model, consists of three principal steps:

1. **HydroWizard Installation:** The framework can be readily installed using the Python package manager:

```
pip install hydrowizard
```

2. **Model Configuration Acquisition:** The model configuration file (`config.yaml`) along with any referenced auxiliary data files (e.g., bathymetry data for reservoirs) should be obtained by researchers from the provided repository or relevant appendices. These files should be saved in a local directory (e.g., `model/`).

3. **Execution of HydroWizard CLI Commands:** The optimization and simulation processes are encapsulated in two primary commands:

   ```
   hw-optimization <options> -c model/config.yaml
   hw-simulation <options> -c model/config.yaml
   ```

   These commands are designed to be executed on the model configuration file, with various options available to customize the processes as needed.

The `hw-optimization` and `hw-simulation` commands serve as high-level interfaces to the complex underlying processes of multi-objective optimization and water resource system simulation, respectively. The specific command options and a comprehensive explanation of their usage are detailed in Appendix C. For our Omo-Gibe case study, the complete model configuration file (`config.yaml`) and the the structure of the bathymetry data used for the Gibe III and Koysha reservoirs is provided in Appendix D.

This standardized approach to result replication offers several key advantages:

- It facilitates independent verification of research findings, a cornerstone of scientific integrity.
- It enables other researchers to build upon existing models, promoting collaborative advancement in the field.
- It allows for analyses and exploration of alternative scenarios using the same baseline model.
- It enhances the transparency of the modeling process, making complex water resource systems more accessible to a broader audience.

By adopting this replication methodology, we not only ensure the reproducibility of our specific results but also establish a framework for transparent and verifiable research in water resource system modeling. This approach aligns with current best practices in scientific research, promoting open science and facilitating the cumulative growth of knowledge in the field.

This standardized approach not only enhances reproducibility but also significantly streamlines the modeling process for researchers and policy analysts. By utilizing the HydroWizard framework, modelers can focus their efforts primarily on two critical aspects: creating an accurate and comprehensive model configuration file in a human-readable format, and conducting in-depth analysis of the resultant data. The framework abstracts away the complex computational processes that lie between these steps, including the setup of appropriate neural networks for policy representation, execution of evolutionary algorithms for multi-objective optimization, and management of simulation processes. Notably, HydroWizard employs advanced algorithms based on Directed Acyclic Graphs (DAGs) specifically designed and implemented to handle the enforcement of mass balance constraints, satisfaction of system-specific constraints, and simulation processes for any arbitrary water resource system. This sophisticated approach ensures robust and efficient modeling regardless of the system's complexity. The framework's abstraction allows domain experts to concentrate on the hydrological and policy aspects of their work, rather than on the intricacies of computational implementation. Consequently, HydroWizard not only facilitates result replication but also democratizes advanced water resource system modeling, making it more accessible to a broader range of researchers and practitioners in the field, while maintaining the highest standards of computational rigor and flexibility.

5

Results

5.1. Evaluating Convergence of Evolutionary Algorithm

As discussed in the last chapter, we employ the hypervolume indicator to assess the convergence of the NSGA-III algorithm. For each generation of the algorithm, we compute the hypervolume and plot it against the number of function evaluations (NFEs). This process is repeated for multiple algorithm runs, each initialized with a different random seed, to assess the consistency and robustness of the convergence behavior.

The convergence plots in Figure 5.1 reveal several noteworthy features. First, the consistent increase in hypervolume across all runs indicates that the algorithm effectively explores the solution space and improves the quality of solutions over time. Second, the convergence of hypervolume values across different random seeds to similar final values demonstrates the robustness of the NSGA-III algorithm, suggesting that it reliably identifies the same Pareto front regardless of initial conditions.

The logarithmic scale of the main plot highlights the rapid initial improvement in solution quality, followed by a more gradual refinement phase. The linear scale inset provides a complementary view, emphasizing the absolute magnitude of improvements and the final convergence state. Together, these plots offer a comprehensive assessment of the convergence behavior of the NSGA-III algorithm, ensuring that the Pareto front is reliably identified and well-represented by the solutions obtained.

The convergence analysis is crucial for validating the reliability of our optimization results. It ensures that the Pareto-optimal solutions we obtain are not artifacts of premature convergence or insufficient solution space exploration. This robust convergence across multiple runs provides confidence in the trade-off analysis that follows.

5.2. Pareto Front Analysis

The Pareto front consists of non-dominated solutions that illustrate the trade-offs between conflicting objectives. These solutions are deemed optimal because improving one objective would result in the deterioration of another. This array of options allows decision-makers to choose solutions aligned with their specific preferences and priorities.

Figure 5.2 presents a parallel plot visualization of the Pareto front for the Lower Omo-Gibe River Basin. This visualization technique offers several insights into the nature of the trade-offs in our system. Each line crossing the three parallel axes represents a unique solution, with its position on each axis indicating its performance in that objective. The density of lines in different regions of the plot reveals the distribution of solutions across the objective space.

The best solutions for each individual objective, highlighted in different colors, demonstrate the extremes of the trade-off space. The red line, representing the best hydropower solution, shows high performance on the hydropower axis but poor performance on the other two. Conversely, the blue and green lines, representing the best irrigation and environmental solutions, respectively, show the opposite trend.

The scaling of the axes in the parallel plot can influence the interpretation of the results. In Figure 5.2, the axes are scaled to emphasize the differences in objective values. To provide a different perspective, a mini parallel plot in the top right corner of the figure shows the same data using a min-max value

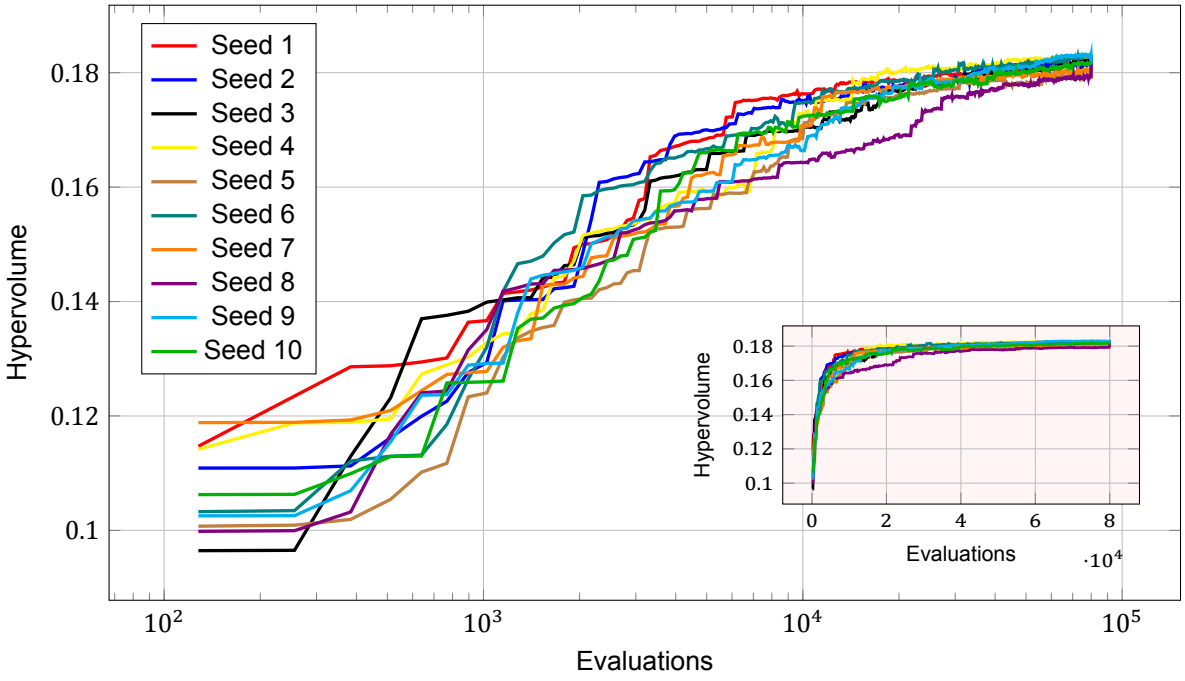


Figure 5.1: Hypervolume vs NFEs for different random seeds to check convergence of MOEA (NSGA-III) on a logarithmic scale. The inset plot in the bottom right shows the convergence on a linear scale.

scale. This alternative scaling allows for a comparative analysis of the solutions and provides additional insights into the trade-offs between the objectives.

5.3. Analysis of Release Decision Rules

The release decision rules, derived from the Radial Basis Function (RBF) network, form the core of our operational policy. These rules determine how water is released from the reservoirs based on current storage levels and the time of year.

In the context of the Lower Omo-Gibe River Basin, the RBF network is employed to approximate the release decisions for the Gibe-III and Koysha reservoirs, as well as the canals to the Kuraz sugar plantations. The RBF network takes the initial storage values of the two reservoirs and the cycle number as inputs, providing the release decisions as outputs. Specifically, the RBF network used in this study has three inputs and provides three outputs. By fixing the cycle number to a specific value, we can visualize the RBF network using 3D plots.

Figure 5.3 presents a novel visualization of these decision rules for the optimal hydropower policy.

This visualization offers several remarkable insights into the operation of our water management system. First, the three-dimensional nature of the plots allows us to see how release decisions vary with both reservoir storage levels simultaneously, revealing complex interactions between the two reservoirs. Second, the ability to animate through different months shows how the release strategy adapts to seasonal water availability and demand variations.

The figure makes the RBF network more interpretable by visually representing the release decision rules. This visualization bridges the gap between the abstract mathematical model and the practical operational decisions that water managers must make. Additionally, it helps the researcher validate the model by visually inspecting the network structure, parameter values, and rule curves, ensuring that they align with the system's expected behavior.

The animated nature of the visualization allows compressing the information from 12 different RBF networks into a single figure, making it easier to convey the intuition behind the neural networks' universal approximation capabilities. This visualization can be used to explain the model to stakeholders and decision-makers, helping them understand how a well-parameterized RBF network can capture complex relationships in the data and provide actionable insights for reservoir operations.

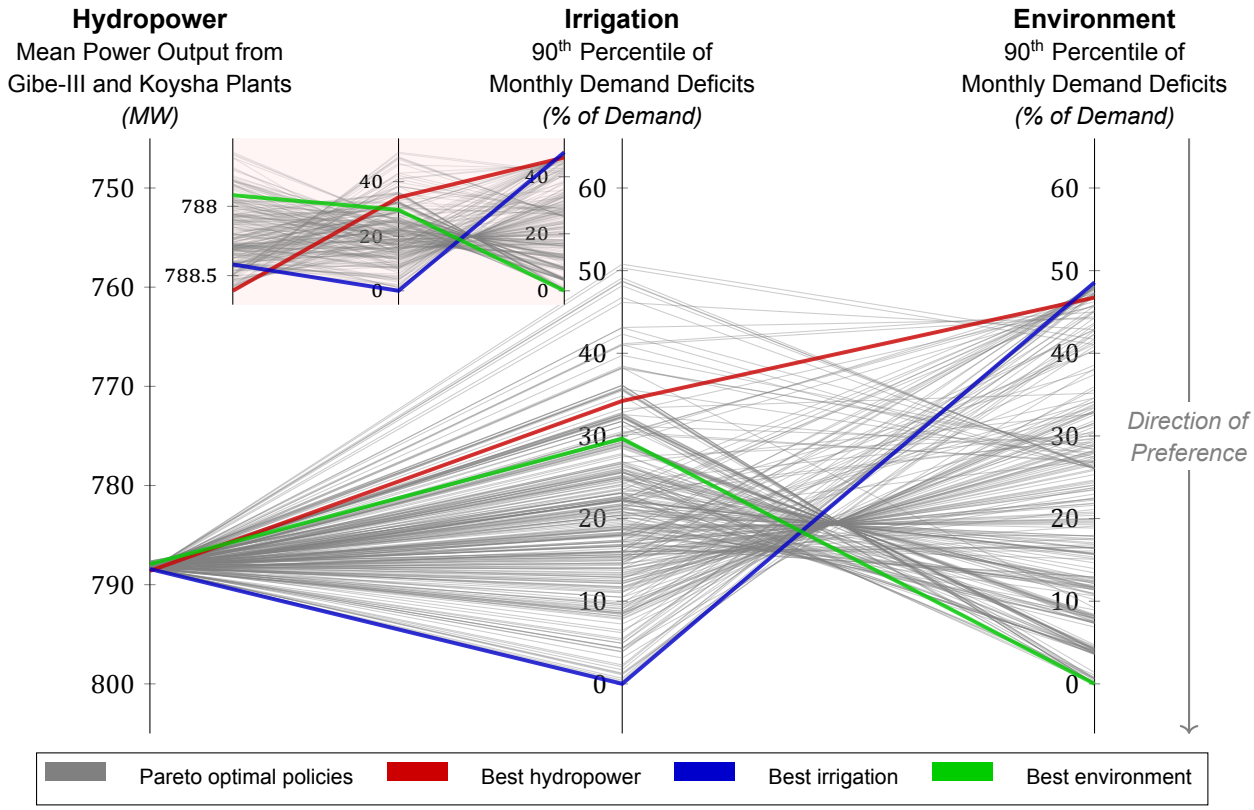


Figure 5.2: Parallel Plot for the solutions in the Pareto Front for the Lower Omo-Gibe River Basin. The score for each objective is computed for the whole simulation horizon. The axis for each objective has been scaled to highlight significant differences in the values. The mini parallel plot in the top right corner shows the same data on a min-max value scale.

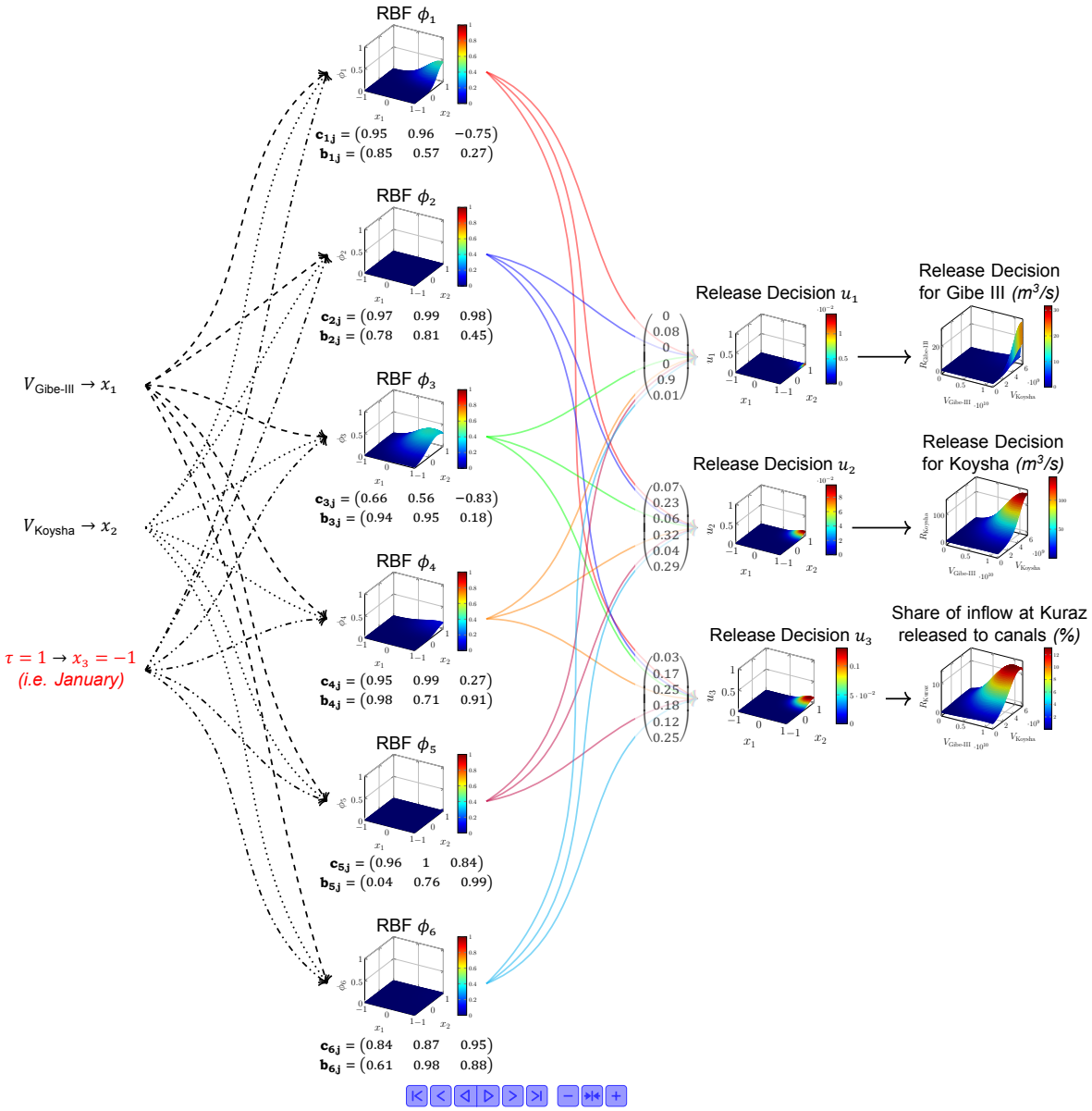


Figure 5.3: The RBF Network parameters (i.e., values for the centers, betas, and weights) in this example correspond to the Best Hydropower policy. The network takes three inputs and provides three outputs. The first two inputs are the initial storage values for Gibe-III and Koysha reservoirs, and the third input is the cycle number τ, which represents the month of the year. The (normalized) values for the first two inputs are plotted on the x and y axes, respectively. The value for τ can be varied from 1 to 12 to get the RBF Network for the corresponding month by moving the animation controls. This allows us to visualize the RBF Network using 3D plots as the input space has been reduced to 2 dimensions, with the third dimension represented by time using animation controls. The rule curves for each of the three release decisions are calculated by taking the weighted sum of the output of the six RBFs. The rule curves can be used to make the release decisions by getting the z-value (denoted by color as well as the height of the surface) for any given initial storage values for each of the two reservoirs (i.e., $V_{\text{Gibe-III}}$ and V_{Koysha}) plotted on the x and y axes. The rule curves add interpretability to the seemingly black-box model of the RBF Network. *Please use Adobe Reader, Foxit Reader, or Okular to see the above figure with animation.*

5.4. Policy Comparison and Model Validation

To provide a comprehensive understanding of how different policies affect system behavior, we compare the rule curves for the optimal hydropower, irrigation, and environment policies. As with the previous visualization, we implement the RBF network independent of the HydroWizard Python package using just the parameter values of the Pareto-optimal solutions. This allows us to visually arrive at the release decisions for each policy that can then be compared with the simulation results obtained from the HydroWizard package.

Figure 5.4 compares rule curves across all months of the year. The animation transitions from one month to the next, allowing us to observe the changes in the release decisions over the year and concisely present a large amount of information in a single figure.

This comparative visualization reveals several noteworthy features. First, we observe distinct differences in the release patterns across the three policies, reflecting their priorities. The best environment policy generally favors higher releases to maximize flow in the main channel, especially during drier months, compared to the other two policies.

Second, the seasonal variation in release rules is evident across all policies, but the magnitude and timing of these variations differ. This reflects the complex interplay between seasonal water availability, varying demands, and the specific objectives of each policy.

The underlying objectives of each policy can explain these differences. The hydropower policy prioritizes maintaining a high hydraulic head and consistent outflow for power generation. In contrast, the irrigation policy aims to store water during wet periods for use during dry spells, while the environment policy seeks to mimic natural flow patterns to support ecosystem health.

This visualization is particularly relevant for water managers and stakeholders. It provides an intuitive understanding of how the system will operate under different conditions, allowing for better planning and potentially increasing trust in the automated decision-making process. Moreover, it bridges the gap between the mathematical complexity of the RBF network and the practical needs of reservoir operators who are accustomed to working with rule curves.

5.5. System State Evaluation

To validate our model and gain deeper insights into system behavior, we analyze the state transitions of the system under different policies.

Figure 5.5 presents a comparison of system state graphs, as introduced in the previous chapter, for the best hydropower, irrigation, and environment policies across selected intervals in the simulation.

These state graphs provide a comprehensive snapshot of the system's behavior at specific time points. They show how water flows through the system, how reservoirs are managed, and how different demands are met under each policy.

These state graphs are particularly relevant for model validation and system understanding. They allow us to verify that mass balance is maintained throughout the system and that all components behave as expected under different management regimes. Moreover, they provide insights into how local decisions (e.g., reservoir releases) impact downstream conditions, which is crucial for integrated water resources management.

5.6. Physical Implications of the Policies

To understand the practical impacts of different management policies, we analyze key physical variables of the system over time.

Figure 5.6 illustrates the year-wise release rates, head values, and power generation for the Gibe-III and Koysha reservoirs under different policies.

This figure reveals several notable patterns. First, we observe that the release rates vary considerably over the months. Second, there is only a subtle difference between policies that optimize for different objectives.

The time-varying nature of the release is due to considerable variations in inflows and demands over the year. The release rates are highest during the wet season when inflows are high, and demands are low and lowest during the dry season when inflows are low, and demands are high. The differences between the policies do not seem significant because of the scale of the plots. However, the minor visible differences significantly impact the deficit minimization, as will be shown in the plots related to

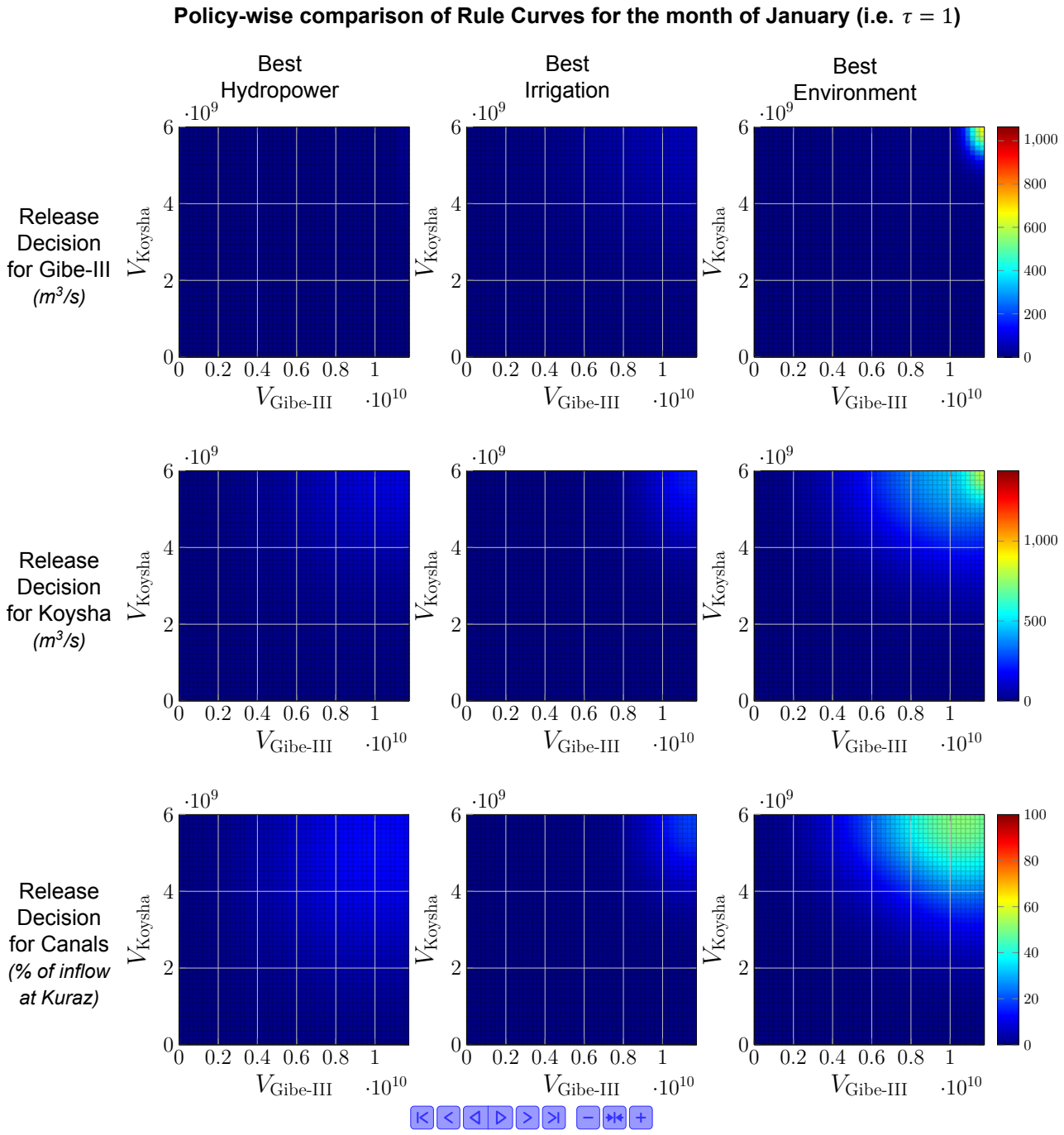


Figure 5.4: Policy-wise comparison of Rule Curves obtained from the RBF network for all three release decisions for each month. The transition from one month to the next is animated to show the changes in the rule curves over the year. The animation can be paused, rewound, and fast-forwarded using the controls provided to study the changes in the rule curves over the year and analyze the policy-wise comparison for any given month. *Please use Adobe Reader, Foxit Reader, or Okular to see the above figure with animation.*

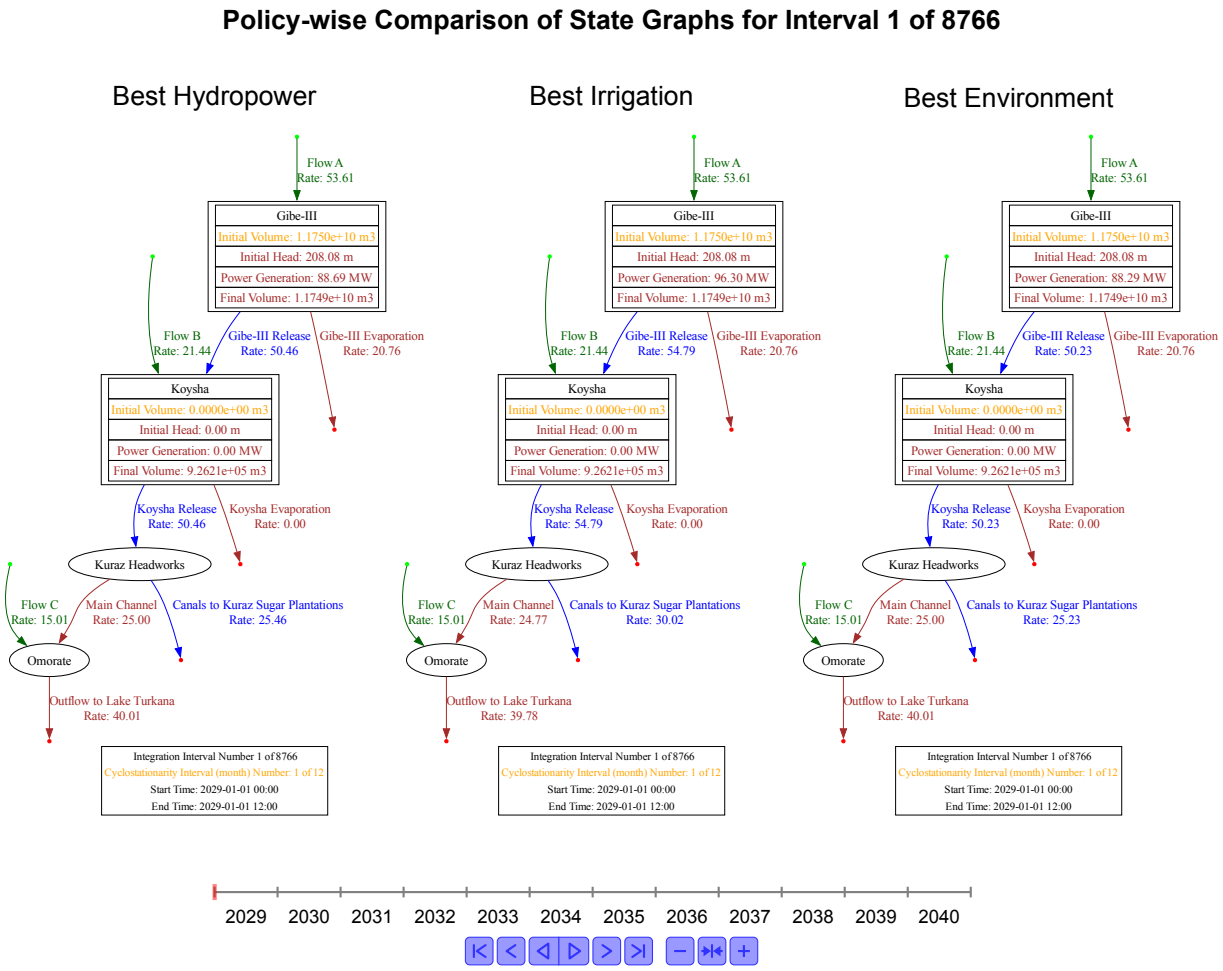


Figure 5.5: Policy-wise comparison of system state graphs for selected intervals in three different model simulations that make release decisions using Best Hydropower, Best Irrigation, and Best Environment policy, respectively. The state graphs show the flow rates in the system for that interval. The green flows are external inflows, the orange variables are the inputs for the RBF network (or the corresponding rule curves), the blue flows are the release decisions derived from the outputs of the RBF network (or the rule curves), and the red flows and state variables are derived using system relationships. *Please use Adobe Reader, Foxit Reader, or Okular to see the above figure with animation.*

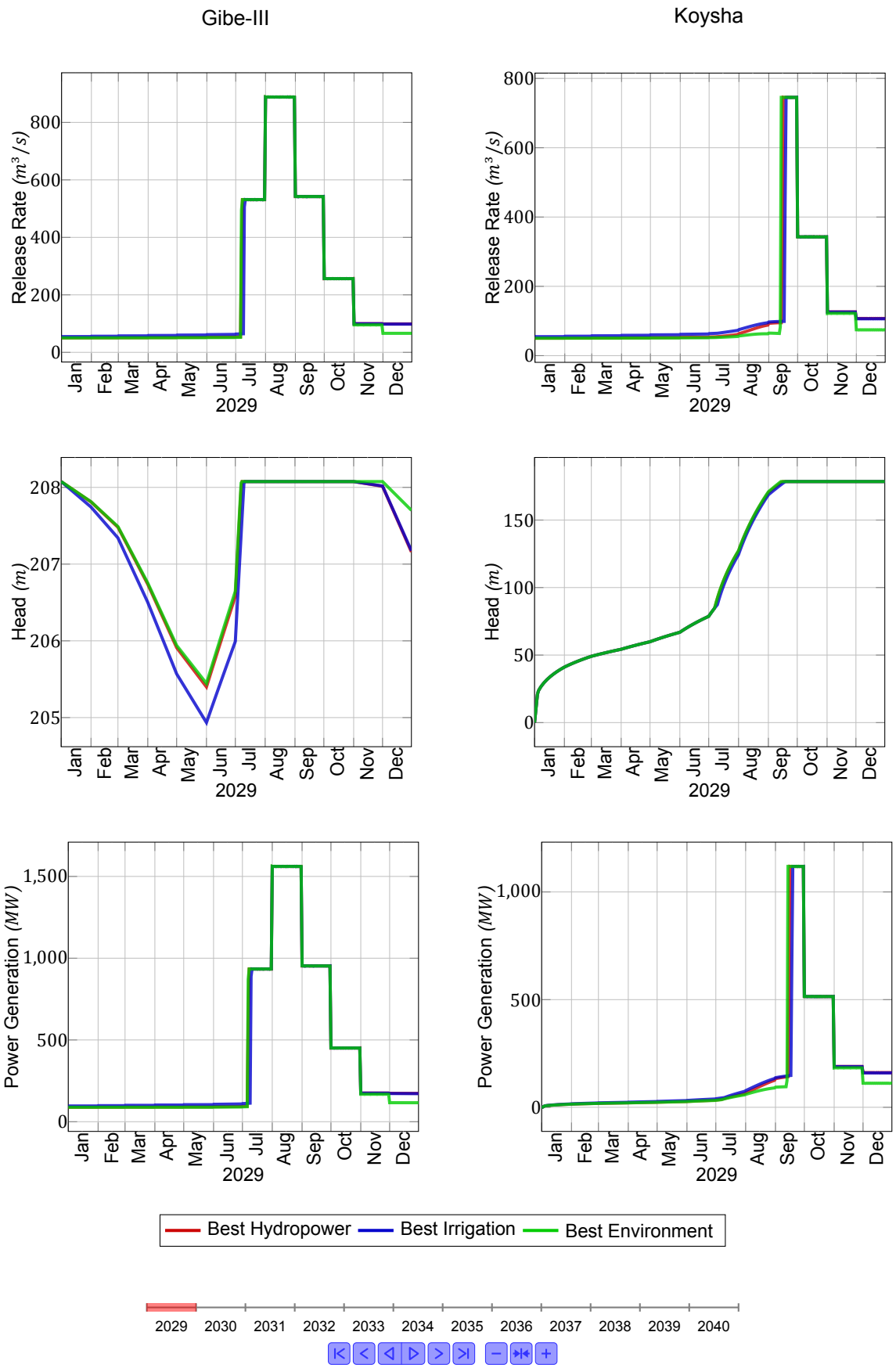


Figure 5.6: Release Rate, Head and Power Generation for Gibe-III and Koysha reservoirs under different management policies.

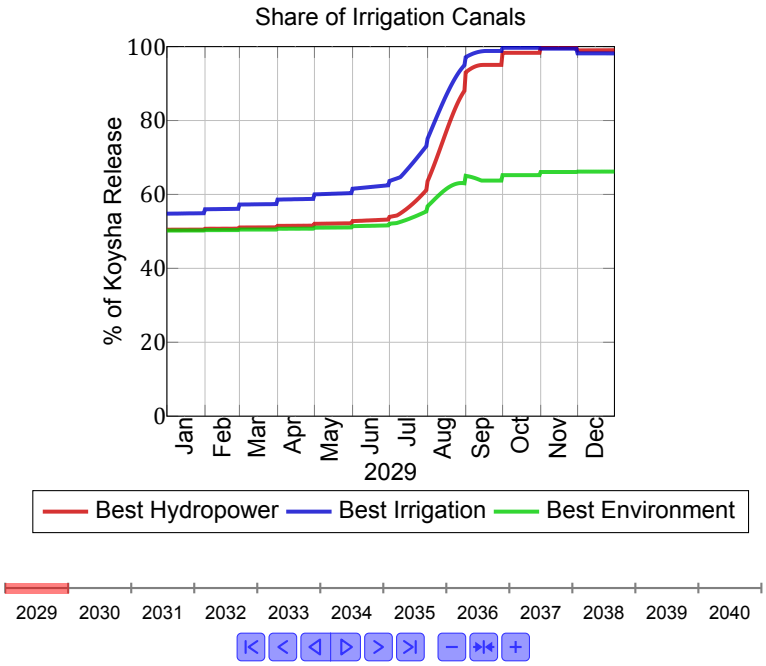


Figure 5.7: Share of Irrigation Canals in Koysha Release under different management policies.

demand deficits.

These results may not be surprising to water managers, as they understand that minor changes in release rates can significantly impact downstream water availability and system performance. However, this visualization can help other stakeholders understand the complexities of water management and appreciate that there is only a tiny wiggle room for policy changes in real-world systems where several physical and operational constraints are at play.

Figure 5.7 focuses on water allocation to irrigation, a key concern in the Lower Omo-Gibe Basin.

This figure illustrates the percentage of Koysha's release allocated to irrigation canals over time. A striking feature is the significant difference in allocation patterns across policies. The irrigation policy, as expected, consistently allocates a higher percentage of flow to irrigation, especially during relatively drier months and high water demand (i.e., October to December).

The hydropower policy, interestingly enough, plays a balancing act between irrigation and power generation,. On the other hand, the environment policy maintains a more consistent allocation to irrigation, reflecting its focus on mimicking natural flow patterns.

These differences stem from the distinct objectives of each policy. This visualization is particularly relevant for agricultural planning and water allocation negotiations in the basin.

Figure 5.8 presents the surface area and evaporation rates for the Gibe-III and Koysha reservoirs under different policies.

This figure is strongly correlated with the figure 5.6. The surface area is directly related to head values.

Interestingly, even though the surface area of the reservoirs is lowest during the dry season, the evaporation rates are highest during this period. This is because the dry season is characterized by high temperatures and low humidity, leading to increased evaporation rates.

This visualization is particularly relevant for making informed decisions, as evaporation rates can impact the overall water balance of the system.

Figure 5.9 focuses on the downstream impacts of different management policies, particularly on irrigation and environmental flows.

This figure comprehensively shows how different policies affect water availability for irrigation and environmental needs. In addition to simulated mean flow rates for different policies, it shows the demand rate for irrigation and environmental flows (in black color). The difference between the demand rate and the mean flow rate is the deficit, shown in the middle row. The deficit percentage is shown in the bottom row.

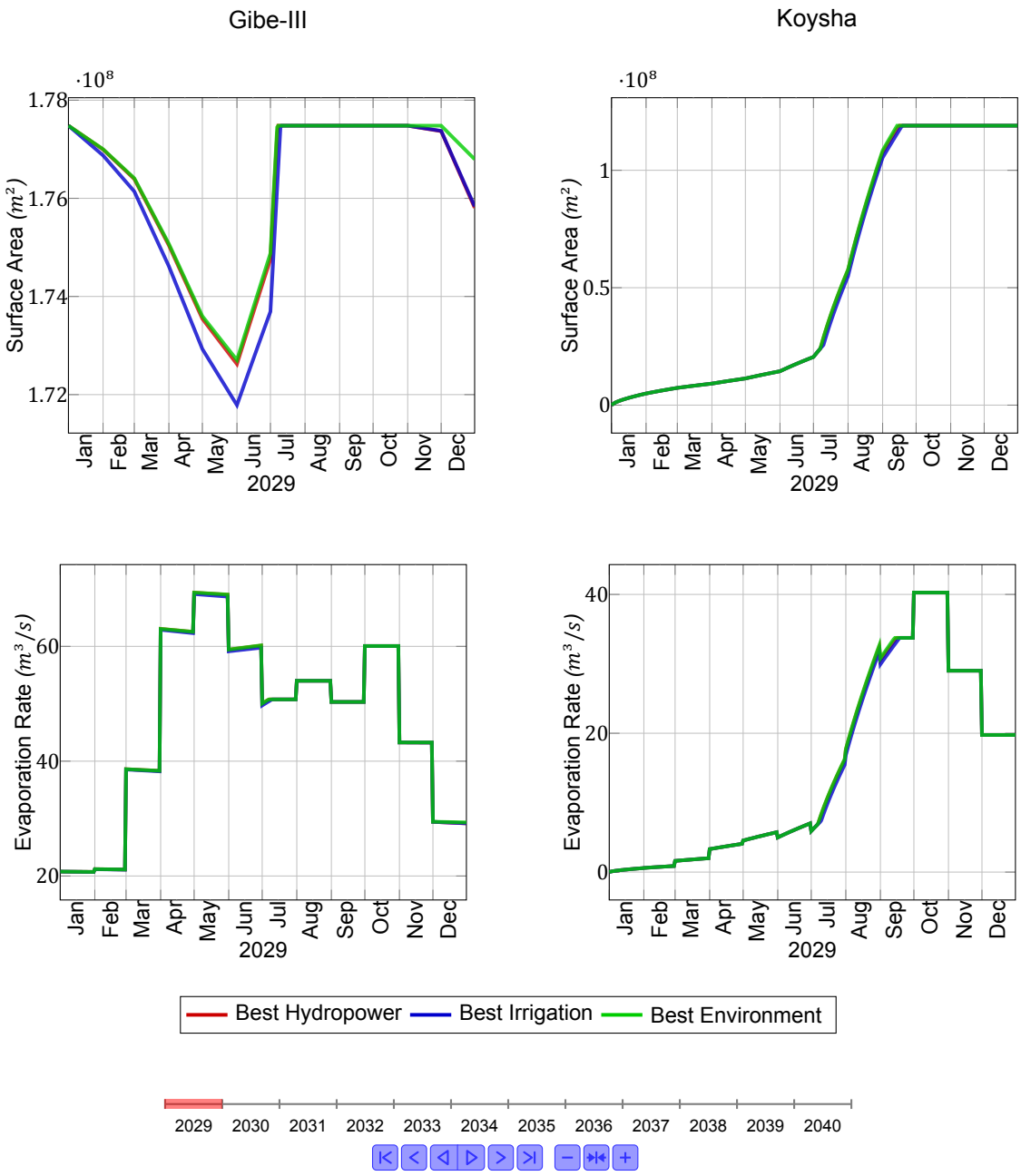


Figure 5.8: Surface Area and Evaporation Rate for Gibe-III and Koysha reservoirs under different management policies.

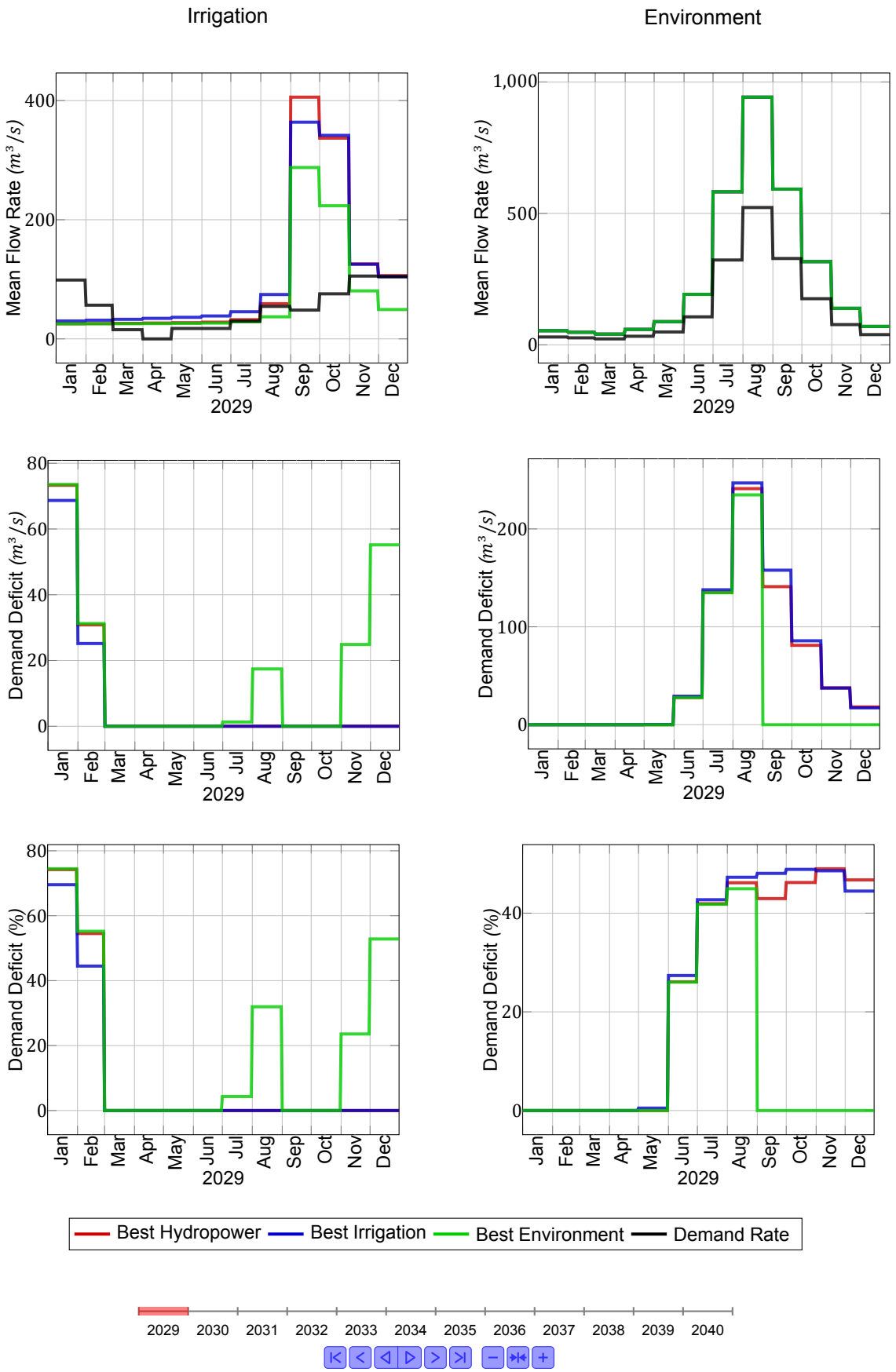


Figure 5.9: Mean Flow Rate, Deficit and Deficit Percentage for Irrigation and Environment under different management policies.

A striking feature is the apparent trade-off between meeting irrigation demands and maintaining environmental flows.

We observe that the irrigation policy results in higher mean flow rates and lower deficits for irrigation but at the cost of more frequent and severe deficits in environmental flows, especially from September to October. Conversely, the environment policy maintains more consistent environmental flows but leads to higher irrigation deficits, particularly during November to February.

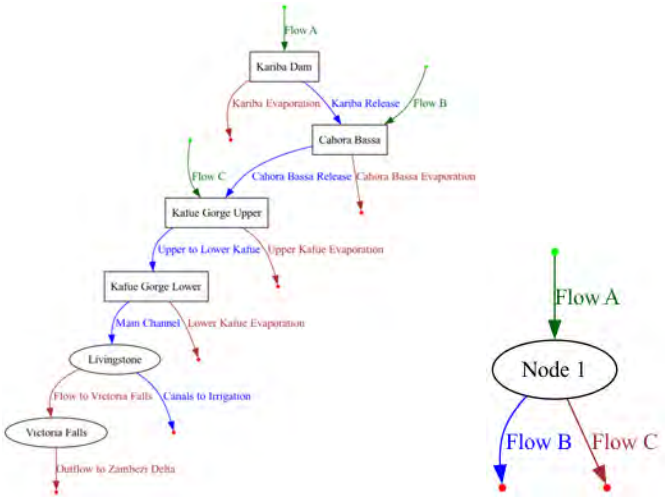
The competing nature of irrigation and environmental water demands can explain these patterns. Water allocated to irrigation is unavailable for environmental flows, and vice versa. The hydropower policy, focused primarily on reservoir management for power generation, often results in intermediate outcomes for irrigation and environmental flows.

This analysis clearly illustrates the challenges of balancing human water needs (represented by irrigation) with ecosystem requirements (represented by environmental flows).

5.7. Broader Applicability of the HydroWizard Framework

To demonstrate the versatility and scalability of our approach, we applied the HydroWizard framework to other river basins of varying complexity.

Figure 5.10 presents the Directed Acyclic Graph (DAG) representations of the Zambezi River Basin and a minimal basin, as parsed from their specification files by the HydroWizard package.



(a) Zambezi River Basin (b) Minimal Basin

Figure 5.10: DAG representation of the Zambezi River Basin and a minimal basin.

This figure showcases the flexibility of our modeling framework in handling systems of different scales and complexities. The Zambezi basin model (Figure 5.10a) represents a real-world, complex river system with multiple reservoirs and demand nodes. The minimal basin (Figure 5.10b) demonstrates the framework's ability to handle a simple hypothetical system.

A remarkable feature of these representations is how they capture the essential structure of each basin, regardless of its complexity. The nodes represent various system elements (e.g., reservoirs, junctions, demand sites), while the edges represent water flows between these elements.

The ability of the framework to handle this range of system complexities is due to its modular design and the use of a generalized graph-based representation. This approach allows for the easy addition or removal of system components without requiring changes to the core optimization and simulation algorithms.

The relevance of this demonstration extends beyond the specific basins shown. It validates the HydroWizard framework as a versatile water resource management tool across a wide range of contexts. This flexibility is crucial for addressing the diverse challenges faced in different river basins worldwide, from small, simple systems to large, complex ones.

6

Discussion

This chapter synthesizes our findings, addresses the research question posed at the outset of this study, and explores the broader implications of our work. We examine how the HydroWizard framework and its application to the Lower Omo-Gibe River Basin contribute to our understanding of water resource management and offer insights for policymakers and researchers alike. The discussion is structured to reflect on the methodological innovations, technical advancements, and practical applications of our research, culminating in an analysis of its policy implications and contributions to the field of water resource management.

6.1. Addressing the Research Question

An overarching research question guided our study:

> *"How can we transparently and efficiently model a water resource system and search for optimal policies for competing water uses, considering uncertainties from evolving development projects and data scarcity, as exemplified by the Lower Omo Basin?"*

The HydroWizard framework, developed and applied in this study, directly addresses this question by providing a unified, transparent approach to water resource modeling and optimization. This framework enables efficient exploration of complex water systems even in data-scarce environments, a critical capability for addressing real-world water management challenges.

As demonstrated in the Lower Omo-Gibe Basin case study, the framework's ability to handle competing water uses and evolving development scenarios showcase its potential to tackle multifaceted water management issues. By integrating stakeholder objectives, hydrological constraints, and optimization techniques within a single, cohesive platform, HydroWizard offers a comprehensive solution to the challenge posed by our research question.

Moreover, the framework's emphasis on transparency, achieved through its YAML-based Model Specification Language, addresses an essential aspect of our research question. This transparency facilitates better understanding among stakeholders and enhances the reproducibility of water resource modeling studies, a crucial factor in building trust and credibility in the decision-making process.

6.2. Methodological Innovation: The HydroWizard Framework

The primary methodological innovation of this study is the development of the HydroWizard framework, a general-purpose approach for water resource optimization and simulation. This framework represents a significant advancement in the field, integrating well-established concepts such as the XLRM framework and Evolutionary Multi-Objective Direct Policy Search (EMODPS) into a cohesive, flexible platform for water resource modeling.

At the core of HydroWizard's innovation is its unified approach to complex water resource modeling. The framework streamlines the modeling process by automating the application of XLRM and EMODPS principles to diverse water systems, reducing the time and expertise required to develop sophisticated

water resource models. This automation enhances efficiency and minimizes the potential for human error in model construction.

The framework's transparent, YAML-based Model Specification Language represents a paradigm shift in how water resource models are defined and communicated. This human-readable format bridges the gap between technical experts and stakeholders, facilitating better understanding and collaboration in the modeling process. The clarity and accessibility of the model specifications enhance the reproducibility of studies, a critical factor in scientific research and policy development.

Central to HydroWizard's design is the clear separation between the Model Specification Language and the Model Execution Engine. This separation represents a fundamental shift in how water resource models can be constructed and evolved. Users can define complex water systems using a consistent, human-readable syntax without needing to grapple with the intricacies of the underlying computational algorithms. Meanwhile, developers can enhance or replace algorithms in the Model Execution Engine without requiring users to modify their existing model specifications.

The implications of this modular architecture are far-reaching. It fosters a more collaborative and inclusive approach to water resource modeling, where domain experts can focus on accurately representing water systems. At the same time, computational scientists can work on improving algorithms and efficiency. This division of labor has the potential to accelerate innovation in water resource management by allowing specialists to contribute their expertise more effectively.

A key feature of the HydroWizard framework is its robust methodology for results replication. This approach, detailed in the Methods chapter, ensures that all studies using HydroWizard can be independently verified and built upon by other researchers. By providing a standardized process for sharing model configurations and executing simulations, HydroWizard sets a new standard for reproducibility in water resource modeling.

Furthermore, the Model Execution Engine's use of a Directed Acyclic Graph (DAG) representation for water systems is a key methodological advancement. This approach enhances computational efficiency and provides a robust framework for ensuring mass balance and enforcing physical constraints within the model. The DAG structure allows for efficient traversal of the water system, facilitating rapid simulation and optimization of complex water networks.

The methodological innovations embodied in HydroWizard have significant implications for water resource management. By lowering the barriers to sophisticated modeling and optimization, the framework democratizes access to advanced water management techniques. This could enhance decision-making processes, particularly in regions with limited resources or technical expertise.

Moreover, the framework's flexibility and extensibility position it as a versatile platform capable of adapting to emerging challenges in water management. This adaptability will be crucial as we face increasingly complex water resource issues, from climate change impacts to growing urban demands. The framework can evolve to incorporate new modeling techniques, data sources, or management strategies while maintaining a consistent and familiar user interface.

6.3. Technical Advancements

The development of the HydroWizard framework has led to several significant technical advancements that enhance our ability to model, optimize, and communicate complex water resource systems. These advancements address key challenges in the field of water resource management, from computational efficiency to the visualization of complex data. This section discusses three primary areas of technical innovation: visualization techniques, computational efficiency, and the framework's implementation and extensibility.

6.3.1. Visualization Techniques

Our novel visualization techniques, particularly the animated rule curves and state graphs, significantly enhance the interpretability of complex water management strategies. The animated rule curves offer an intuitive understanding of how reservoir operations evolve in response to changing conditions, transforming static data into a straightforward, engaging visual narrative of water management decisions. This dynamic representation allows stakeholders to grasp the temporal aspects of water management policies, a crucial factor in understanding system behavior over time.

Complementing these are our state graphs, which provide comprehensive snapshots of the entire water system at specific points in time. These graphs offer a holistic view of system dynamics by

capturing the interplay between various system components, allowing stakeholders to understand how components interact within the broader water management context.

These visualization techniques serve as a crucial bridge between sophisticated algorithms and practical decision-making, transforming complex, multi-dimensional data into accessible, intuitive visual representations. By making the results of complex optimizations more accessible, these visualizations play a vital role in democratizing water resource management, enabling more inclusive, informed discussions about water allocation strategies.

6.3.2. Computational Efficiency

The computational efficiency achieved by the Python implementation of HydroWizard is a key feature that significantly enhances its applicability to complex water resource systems. Our implementation's built-in support for multiprocessing is essential for handling large hydrological models efficiently, an increasingly important capability as we model more complex and interconnected water systems.

Our computational setup achieved a CPU utilization of 96.15% with a population size of 128, 625 generations, and 64 processes. This high utilization across multiple cores demonstrates the scalability of the HydroWizard implementation, allowing it to leverage modern multi-core processors effectively. The ability to parallelize computations reduces the time required for complex simulations and optimizations and enables the exploration of larger solution spaces, potentially leading to better-optimized water management strategies.

By incorporating multiprocessing support out of the box, the HydroWizard implementation ensures that users can readily apply it to large-scale, real-world water systems without implementing parallelization strategies. This efficiency allows for rapid exploration of different scenarios and policy options, a crucial capability in the face of changing climate and socio-economic conditions.

6.3.3. Implementation and Extensibility

To demonstrate the technical feasibility of the HydroWizard framework, we developed an open-source Python implementation. This proof-of-concept makes advanced modeling techniques accessible to a broader range of users and applicable to various water systems. The implementation's modular design allows for easy extension and modification, enabling researchers and practitioners to adapt the framework to their specific needs.

The framework's design and documentation are intentionally implementation-agnostic, allowing for potential realization in various programming languages beyond our provided Python implementation. This approach ensures that the core principles and structure of HydroWizard can be leveraged across different technological ecosystems, broadening its potential impact and longevity in the field.

The framework's extensibility is particularly important in the rapidly evolving field of water resource management. As new modeling techniques, optimization algorithms, or data sources become available, they can be integrated into the framework without requiring a complete overhaul of existing models.

6.4. Application to the Lower Omo-Gibe Basin

Our application of HydroWizard to the Lower Omo-Gibe Basin demonstrates the framework's capability to handle complex, real-world water management challenges. We identified three primary groups in the basin through stakeholder analysis: Ethiopian Electric Power (EEP), the Ethiopian Sugar Industry Group (ESIG), and local communities. Each group's objectives were quantified and incorporated into our model as key performance metrics: mean power output from Gibe-III and Koysha plants, 90th percentile of monthly irrigation demand deficits, and 90th percentile of monthly environmental flow deficits.

Our model configuration captures the basin's essential dynamics despite data scarcity challenges. This involved carefully selecting and estimating key parameters, integrating available data, and developing proxy measures where direct data was unavailable. The resulting model adequately represents the basin's hydrology, infrastructure, and water demands.

The results of our analysis reveal significant trade-offs among the three primary objectives. The 'Best Irrigation' policy eliminates irrigation deficits but increases environmental flow deficits by up to 48% in critical months. Conversely, the 'Best Environmental Flow' policy maintains optimal river flows but can increase irrigation deficits up to 51% during peak demand periods. Interestingly, we observed

surprisingly stable mean power generation across different Pareto-optimal policies, with only a 0.9 MW difference between the best and worst-performing policies.

These findings highlight the complex compromises inherent in water resource management and demonstrate the framework's ability to provide decision-makers with pareto-approximate policy choices. By quantifying these trade-offs, our model provides a solid foundation for informed decision-making in the basin.

6.5. Policy Implications

The insights gained from our analysis have significant implications for water resource management in the Lower Omo-Gibe Basin and beyond. Quantifying trade-offs between different water uses provides a solid foundation for informed policy discussions. Policymakers now have a clearer understanding of the consequences of prioritizing one objective over others, allowing for more strategic and balanced decision-making.

The resilience observed in hydropower generation across different management strategies suggests that energy production need not be the sole driver of water allocation decisions. This finding could encourage a more holistic approach to water management, giving greater weight to ecological and social considerations without significantly compromising energy security. It challenges the often-assumed strict trade-off between energy production and other water uses, potentially opening up new avenues for sustainable water management.

However, our results also highlight the need for adaptive management strategies. The temporal dynamics revealed by our analysis, particularly the seasonal variations in water availability and demand, underscore the importance of flexible policies that can respond to changing conditions. This might involve developing more sophisticated real-time control systems or implementing seasonal variations in management strategies.

The clear trade-offs between irrigation and environmental flows point to the need for innovative approaches to agricultural water use. Investments in more efficient irrigation technologies or shifts to less water-intensive crops could alleviate some of the pressure on the basin's water resources. Similarly, exploring alternative livelihoods or compensation mechanisms for ecosystem services could help balance economic development with ecological preservation.

Furthermore, the framework's ability to handle data scarcity and provide meaningful insights, even with limited information, has important implications for water management in developing regions. It demonstrates that sophisticated modeling and optimization techniques can be applied even in contexts where comprehensive data is unavailable, potentially improving water management practices in areas where advanced modeling approaches have historically been underserved.

6.6. Limitations and Future Research Directions

While this study makes significant contributions to water resource management, there are exciting opportunities for future research that could further enhance the capabilities of the HydroWizard framework and its applications.

One promising avenue is the integration of comprehensive uncertainty analysis. Future work could assess the robustness of proposed policies under different scenarios by incorporating variability in climate projections, socio-economic scenarios, and other external factors. This would give policymakers even more confidence in the framework's recommendations and help develop adaptive strategies for a range of possible futures.

Data quality and availability, while challenging in many contexts, present opportunities for innovation. Future research could explore advanced data assimilation techniques, remote sensing applications, and citizen science approaches to enhance data collection and management. These efforts could improve model accuracy and contribute to broader initiatives for water resource monitoring and management.

Extending the simulation timeframe in future studies could offer valuable insights into long-term system behavior, particularly in climate change. Incorporating climate projections into the model could help identify robust management strategies under different future scenarios, enhancing the framework's utility for long-term planning and adaptation.

An exciting direction for future research is the exploration of accelerated computing using Graphics Processing Units (GPUs). The parallel processing capabilities of GPUs could significantly enhance the

computational efficiency of HydroWizard, particularly for large-scale simulations and complex optimization problems. GPU acceleration could enable the exploration of larger solution spaces, more detailed simulations, and faster real-time decision support, pushing the boundaries of what is possible in water resource modeling and optimization.

The transboundary nature of many water systems, including the Omo-Gibe Basin's impact on Lake Turkana in Kenya, suggests the potential for broader regional analyses. Future research could extend the model to include downstream impacts and explore cooperative management strategies, contributing to more comprehensive and equitable water resource management across regions.

Further refinement of the framework could involve exploring different percentiles or temporal scales of demand deficits, offering a more nuanced picture of system dynamics. Analyzing system behavior under rare but high-impact scenarios could inform the development of more resilient water management strategies, enhancing the framework's ability to support robust decision-making.

7

Conclusion

This chapter synthesizes the key findings, contributions, and implications of our research on transparent and efficient water resource modeling, directly addressing our vision of ”Modeling Water Resources for Everyone”.

7.1. Summary of Key Contributions

Our research tackles the challenge of making water resource modeling accessible to a broad audience. The HydroWizard framework, developed and applied to the Lower Omo-Gibe River Basin, embodies our commitment to transparent and effective approaches for complex water systems.

By integrating a YAML-based Model Specification Language with an efficient Model Execution Engine, HydroWizard achieves unprecedented transparency and modularity. This innovative approach democratizes complex system modeling, reducing development time from months to hours and making it accessible to a wider range of users. The framework's architecture, separating model specification from execution, enhances reproducibility and facilitates collaborative research, truly making water resource modeling available for everyone.

A key contribution is the integration of cutting-edge optimization techniques, specifically Evolutionary Multi-Objective Direct Policy Search (EMODPS) and Radial Basis Function (RBF) networks. This integration creates a powerful tool for exploring complex water management scenarios, demonstrating effectiveness in handling intricate systems like the Lower Omo Basin.

HydroWizard's implementation as an open-source Python package, freely available on GitHub and PyPI, represents a significant step towards democratizing advanced water resource modeling. This approach not only supports our goal of modeling water resources for everyone but also allows for continuous improvement and adaptation by the wider scientific community, potentially accelerating innovations in water management practices.

Our novel visualization techniques, particularly animated rule curves and state graphs, enhance the interpretability of complex water management strategies. These tools bridge the gap between sophisticated algorithms and practical decision-making, making water resource modeling more transparent and accessible to non-experts.

The application of HydroWizard to the Lower Omo-Gibe Basin yielded nuanced insights into the complex interplay of hydropower generation, irrigation demands, and environmental flow requirements. This real-world application demonstrates the framework's ability to provide transparent and effective approaches for complex systems, quantifying trade-offs between conflicting objectives in a data-scarce environment.

7.2. Implications and Impact

The implications of this work directly support our goal of modeling water resources for everyone. HydroWizard has the potential to significantly impact water resource management practices, particularly in regions facing acute water stress and limited technical resources. Its ability to handle data scarcity and evolving development scenarios makes it especially valuable in developing countries, truly democratizing advanced water resource modeling.

The Lower Omo-Gibe Basin case study showcases the framework's capacity to provide actionable insights in a complex environment. By identifying Pareto-approximate solutions, it offers decision-makers a range of policy options, exemplifying how transparent and effective approaches can be applied to real-world complex systems.

The framework's ability to rapidly generate and analyze multiple scenarios positions it as a valuable tool for adaptive management strategies. This aligns with our goal of providing effective approaches for complex systems, crucial in an era of climate change and rapid socio-economic development.

7.3. Future Directions and Concluding Remarks

Looking forward, we envision further advancements in making water resource modeling accessible to everyone. Integrating HydroWizard with climate change models could provide crucial insights into long-term water security strategies. Incorporating more sophisticated ecological models could deepen our understanding of the relationship between water management decisions and ecosystem health, further enhancing the effectiveness of our approach for complex systems.

In conclusion, this research significantly contributes to realizing our vision of "Modeling Water Resources for Everyone". The HydroWizard framework represents a paradigm shift in approaching the modeling and optimization of complex water systems, offering transparent and effective approaches as demonstrated in the Lower Omo Basin case study. This work bridges the gap between advanced academic research and practical water management. The HydroWizard framework is a testament to the power of interdisciplinary research and open-source development in tackling complex environmental challenges.

As we move forward, we hope the methodologies and insights developed in this study will catalyze more informed, equitable, and sustainable water resource management practices worldwide. By providing a flexible, transparent, and efficient platform for water resource modeling, we equip researchers, policymakers, and practitioners with the tools needed to navigate an uncertain water future. In doing so, this research contributes to the broader goal of ensuring water security in an increasingly water-stressed world, paving the way for more resilient and sustainable water management practices in the years to come.

A

HydroWizard Model Specification Language

A.1. Introduction

The HydroWizard Model Specification Language provides a formal framework for defining water resource systems. It uses YAML (YAML Ain't Markup Language) as its base format, chosen for its simplicity, readability, and flexibility. This appendix serves as a comprehensive reference manual for the language, detailing its structure, components, and usage.

A.2. Basic Structure

A HydroWizard model specification consists of four main sections:

1. Basin Configuration
2. Nodes
3. Flows
4. Objectives

Each section is defined as a top-level key in the YAML document.

A.3. Basin Configuration

The basin configuration section defines the overall parameters of the water resource system.

A.3.1. Syntax

```yaml
basin:
  name: <string>
  simulation_start_time: <YYYY-MM-DD HH:MM>
  simulation_horizon: <integer>
  integration_interval: <integer>
  cyclostationarity_interval: <string>
  optimization_method: <string>
```

A.3.2. Fields

- `name`: A string identifier for the basin.
- `simulation_start_time`: The start time of the simulation in "YYYY-MM-DD HH:MM" format.

- `simulation_horizon`: The duration of the simulation in years.
- `integration_interval`: The time step for integration in hours.
- `cyclostationarity_interval`: The interval for cyclostationarity. Valid options are ”Month”, ”Half Month”, ”Two Month”, ”Quarter”.
- `optimization_method`: The method used for optimization. Currently, only ”MOEA RBF” is supported.

A.4. Nodes

The nodes section defines the discrete points in the water resource system, such as reservoirs, confluences, and measurement stations.

A.4.1. Syntax

```
nodes:
  - name: <string>
    reservoir_node: <boolean>
    initial_volume: <float>
    max_volume: <float>
    min_volume: <float>
    bathymetry: <string>
    evaporation_rate: <float or list or string>
    power_generation_node: <boolean>
    turbine_max_power: <float>
    turbine_efficiency: <float>
    turbine_head: <float>
    turbine_max_flow_rate: <float>
```

A.4.2. Fields

- `name`: A unique identifier for the node.
- `reservoir_node`: Boolean indicating if the node is a reservoir (true) or not (false).
- `initial_volume`: Initial water volume in m³ (for reservoir nodes).
- `max_volume`: Maximum water volume in m³ (for reservoir nodes).
- `min_volume`: Minimum water volume in m³ (for reservoir nodes).
- `bathymetry`: Filename of the CSV file containing bathymetry data (for reservoir nodes).
- `evaporation_rate`: Monthly evaporation rates in mm/month. Can be a single float, a list of 12 values, or a CSV filename.
- `power_generation_node`: Boolean indicating if the node generates hydropower.
- `turbine_max_power`: Maximum power output of the turbine in MW.
- `turbine_efficiency`: Efficiency of the turbine (0-1).
- `turbine_head`: Head of the turbine in meters.
- `turbine_max_flow_rate`: Maximum flow rate through the turbine in m³/s.

A.5. Flows

The flows section defines the water transfers between nodes, including inflows, outflows, and releases.

A.5.1. Syntax

```yaml
flows:
  - name: <string>
    source_node: <string or null>
    target_node: <string or null>
    kind: <string>
    flow_rate: <float or list or string>
    demand_rate: <float or list or string>
    evaporation_flow: <boolean>
```

A.5.2. Fields

- `name`: A unique identifier for the flow.
- `source_node`: The name of the node where the flow originates (null for inflows).
- `target_node`: The name of the node where the flow ends (null for outflows).
- `kind`: The type of flow. Valid options are ”x” (external), ”l” (decision-dependent), or ”r” (system-dependent).
- `flow_rate`: Flow rates for external flows. Can be a single float, a list of values, or a CSV filename.
- `demand_rate`: Demand rates for flows with targets. Can be a single float, a list of values, or a CSV filename.
- `evaporation_flow`: Boolean indicating if the flow represents evaporation.

A.6. Objectives

The objectives section defines the optimization objectives for the water resource system.

A.6.1. Syntax

```yaml
objectives:
  - name: <string>
    kind: <string>
    target_node: <string or list>
    target_flow: <string>
    quantile: <float>
```

A.6.2. Fields

- `name`: A unique identifier for the objective.
- `kind`: The type of objective. Valid options are ”Power Generation Maximization” or ”Monthly Demand Deficit Minimization”.
- `target_node`: The node(s) associated with the objective (for power generation).
- `target_flow`: The flow associated with the objective (for demand deficit).
- `quantile`: The quantile used for demand deficit calculation (0-1).

A.7. Data Input Formats

A.7.1. Bathymetry Data

Bathymetry data should be provided in a CSV file with the following columns:

1. Volume (m³)
2. Surface area (m²)
3. Head (m)

A.7.2. Time Series Data

Time series data (e.g., flow rates, evaporation rates) can be provided in three ways:

1. As a single float value (constant for all time steps)
2. As a list of 12 float values (one for each month)
3. As a CSV file with 12 values (one for each month)

A.8. Example

Here's a simple example of a HydroWizard model specification:

```yaml
basin:
  name: "Example Basin"
  simulation_start_time: "2025-01-01 00:00"
  simulation_horizon: 10
  integration_interval: 24
  cyclostationarity_interval: "Month"
  optimization_method: "MOEA RBF"

nodes:
  - name: "Reservoir A"
    reservoir_node: true
    initial_volume: 1000000000.0
    max_volume: 1500000000.0
    min_volume: 500000000.0
    bathymetry: "Bathymetry_A.csv"
    evaporation_rate: [100, 90, 80, 70, 60, 50, 40, 50, 60, 70, 80, 90]
    power_generation_node: true
    turbine_max_power: 100.0
    turbine_efficiency: 0.9
    turbine_head: 50.0
    turbine_max_flow_rate: 200.0

flows:
  - name: "Inflow A"
    source_node: null
    target_node: "Reservoir A"
    kind: "x"
    flow_rate: "inflow_A.csv"
  - name: "Release A"
    source_node: "Reservoir A"
    target_node: null
    kind: "l"

objectives:
  - name: "Hydropower"
```

```yaml
    kind: "Power Generation Maximization"
    target_node: "Reservoir A"
  - name: "Environmental Flow"
    kind: "Monthly Demand Deficit Minimization"
    target_flow: "Release A"
    quantile: 0.9
```

Listing 1: Example of a HydroWizard model specification in YAML format

This example defines a simple water resource system with one reservoir, an inflow, a release, and two objectives (hydropower generation and environmental flow).

B

HydroWizard Core Algorithms

This appendix presents the detailed algorithms referenced in the Methods chapter. These algorithms form the core of the flow computation process in the HydroWizard framework.

B.1. Algorithm for Computing Min/Max Constraints on L-flows

Algorithm 1 details the process of computing min/max constraints on L-flows derived from dependent flow constraints. This algorithm is crucial for ensuring that the flow rates in the system respect physical and operational constraints. This algorithm performs a backward pass through the DAG, propagating

Algorithm 1 Computation of min/max constraints on L-flows derived from dependent flow constraints

Input: Directed Acyclic Graph (DAG) of the water resource system with min/max constraints for some of the flows
Output: Min/max constraints for L-flows in the DAG accumulated from min/max constraints on dependent flows
Initialize: List of leaf nodes in the DAG
Back Propagation of Constraints:
for each leaf node in the DAG **do**
    Initialize current node as a leaf node
    while current node is not the root node **do**
        Identify parent node
        Check and propagate min/max constraints from the current node to the parent node:
        if current node has min constraint **then**
            Propagate min constraint to the parent node
        end if
        if current node has max constraint **then**
            Propagate max constraint to the parent node
        end if
        Accumulate and/or cancel constraints at the parent node
        Update current node to parent node
    end while
end for

constraints from leaf nodes to root nodes. It ensures that all downstream constraints are considered when determining feasible flow ranges for decision-dependent flows.

B.2. Algorithm for Assigning Flow Rates to L-flows and Remaining R-flows

Algorithm 2 outlines the process of assigning flow rates to L-flows and remaining R-flows. This algorithm is essential for maintaining mass balance and satisfying all constraints in the water resource system.

Algorithm 2 Assignment of flow rates to L-flows and remaining R-flows

Preprocessing Steps:
Assign values for X-flows and decision-independent R-flows
Compute policy function recommendations for L-flows
Compute min/max constraints on L-flows arising from min/max constraints on dependent flows
Compute min/max constraints on L-flows arising from min/max constraints on reservoirs' volume
Flow Rate Assignment:
Perform topological sorting of nodes in the basin
for each node in topologically sorted order **do**
    Calculate the actual release value for L-flows considering, in order of precedence:
        - Min/Max constraints arising from min/max constraints on reservoir storage levels
        - Min/Max constraints arising from min/max constraints on dependent flows
        - Policy function recommendations
    Assign the actual release value as the flow rate for the L-flow
    if node is not a reservoir **then**
        Assign the flow rate for R-flow at the node using the mass balance equation
    end if
end for
Repeat for all nodes in the basin to determine the remaining flow rates

This algorithm performs a forward pass through the topologically sorted DAG to determine actual flow rates. It integrates policy recommendations, physical constraints, and mass balance principles to compute final flow values, demonstrating the complex interplay between management decisions and system dynamics in the water resource model.

C

HydroWizard CLI Guide

This appendix provides comprehensive documentation for the Command-Line Interface (CLI) tools provided by HydroWizard package for multi-objective optimization and water resource system simulation.

C.1. Optimization

The `hw-optimization` command is used to find Pareto-optimal policies for a specified water resource system. It offers various options to customize the optimization process.

C.1.1. Options

- `-c, --config_file` (required): Path to the basin configuration file.
- `-o, --output_dir` (required): Directory to save the output results.
- `-p, --population_size` (required): Population size for the optimization algorithm.
- `-g, --num_generations` (required): Number of generations for the optimization algorithm.
- `-s, --simulation_horizon`: Simulation horizon in years (optional). Overrides the value in the configuration file.
- `-i, --interval_duration`: Integration interval duration in hours (optional). Overrides the value in the configuration file.
- `-n, --n_processes`: Number of processes to use for optimization (default: 8). Should be set equal to the number of CPU cores for optimal performance.
- `-r, --random_seed`: Random seed for MOEA algorithm (optional). Random seed is randomly generated if not provided.
- `-d, --db_logging`: Enable database logging of optimization results (flag). Requires database connection credentials in the configuration file and the database schema to be set up.
- `-f, --initiate_with_pareto_front`: Initialize the optimization with the current Pareto front (flag) from previous optimization runs of the same model configuration in the remote database. Requires database connection credentials in the configuration file.

C.1.2. Example Usage

```
hw-optimization -c model/config.yaml \
    -o optimization-results \
    -p 100 \
    -g 100 \
    -s 10 \
```

```
    -i 24 \
    -n 64 \
    -r 42 \
    -d \
    -f
```

This example runs an optimization for the model config with a population size of 100, for 100 generations, simulating 10 years with 24-hour intervals, using 64 parallel processes, a random seed of 42, enabling database logging, and initializing with the current Pareto front.

C.2. Simulation

The `hw-simulation` command is used to simulate flows and stocks for a given water resource system using a specified policy.

C.2.1. Options

- `-c, -config_file` (required): Path to the basin configuration file.
- `-p, -policy_source` (required): Policy source for the simulation.
- `-o, -output_dir` (required): Directory to save the output results.
- `-s, -simulation_horizon`: Simulation horizon in years (optional). Overrides the value in the configuration file.
- `-i, -interval_duration`: Integration interval duration in hours (optional). Overrides the value in the configuration file.
- `-n, -policy_names`: Names for the policies (optional).
- `-v, -visualize_intervals`: Intervals to visualize the state graph for (optional).
- `-r, -include_intermediate_results`: Include intermediate results in the Pareto front (flag).

C.2.2. Example Usage

```
hw-simulation -c model/config.yaml \
    -p optimization-results/01/ParetoBestX.txt \
    -o simulation-results \
    -s 12 \
    -i 12 \
    -n "Policy1,Policy2,Policy3" \
    -v "1,50,100" \
    -r
```

This example runs a simulation for the Lower Omo basin using the best policies from a previous optimization. It simulates 12 years with 12-hour intervals, names the policies "Policy1", "Policy2", and "Policy3", visualizes state graphs for intervals 1, 50, and 100, and includes intermediate results in the Pareto front.

C.3. Advanced Usage

C.3.1. Policy Source Options for hw-simulation

The `-policy_source` option in `hw-simulation` supports various formats:

- **Random**: Use `random` to generate a random policy.
- **Best from Database**: Use `best_from_db` to select the best policies from the database from the previous optimization results of the same model configuration. Requires database connection credentials in the configuration file.

- **Best from Latest Optimization**: Use `best_from_latest:/path/to/optimization/results` to select the best policies from the most recent optimization of the same basin in the specified directory.
- **Specific File**: Provide a file path to use specific policy parameters.
- **Multiple Files**: Use semicolon-separated file paths, optionally specifying row numbers (e.g., `file1.txt:1,2;file2.txt`).

C.3.2. Visualization in hw-simulation

The `–visualize_intervals` option allows you to specify which intervals to visualize in the state graph. Use comma-separated values to list specific intervals (e.g., `1,50,100`).

C.4. Output

Both commands generate comprehensive output, including:

- Optimization/Simulation results
- Pareto front data
- State variables (flow rates, node volumes, etc.)
- Visualizations (basin graphs, state graphs, etc.)
- Performance metrics

Results are saved in the specified output directory, organized by basin name and timestamp for easy reference and analysis.

D

Model for the Lower Omo River Basin

D.1. Configuration File

Complete model configuration file `config.yaml` used for the Lower Omo River Basin is provided below.

```yaml
basin:
  name: "Lower Omo"
  simulation_start_time: "2029-01-01 00:00"
  simulation_horizon: 12 # years
  integration_interval: 12 # months
  cyclostationarity_interval: "month"
  optimization_method: "MOEA RBF"

nodes:
  - name: "Gibe-III"
    reservoir_node: True
    initial_volume: 11750000000.0 # m^3
    max_volume: 11750000000.0 # m^3
    min_volume: 7000000000.0 # m^3
    bathymetry: "Bathymetry Gibe-III.csv"
    evaporation_rate: # mm/month
      - 30.32 # January
      - 31.04 # February
      - 56.73 # March
      - 93.36 # April
      - 103.59 # May
      - 89.21 # June
      - 74.13 # July
      - 78.86 # August
      - 73.49 # September
      - 87.72 # October
      - 63.18 # November
      - 43.01 # December
    power_generation_node: True
    turbine_max_power: 1870.0 # MW
    turbine_efficiency: 0.9 # 0-1
    turbine_head: 9.0 # m
    turbine_max_flow_rate: 1064.0 # m^3/s
  - name: "Koysha"
    reservoir_node: True
```

```yaml
    initial_volume: 0.0 # m^3
    max_volume: 6000000000.0 # m^3
    min_volume: 3000000000.0 # m^3
    bathymetry: "Bathymetry Koysha.csv"
    evaporation_rate: # mm/month
      - 30.32 # January
      - 31.04 # February
      - 56.73 # March
      - 93.36 # April
      - 103.59 # May
      - 89.21 # June
      - 74.13 # July
      - 78.86 # August
      - 73.49 # September
      - 87.72 # October
      - 63.18 # November
      - 43.01 # December
    power_generation_node: True
    turbine_max_power: 2160.0 # MW
    turbine_efficiency: 0.9 # 0-1
    turbine_head: 8.5 # m
    turbine_max_flow_rate: 1440.0 # m^3/s
  - name: "Kuraz Headworks"
  - name: "Omorate"

flows:
  - name: "Flow A"
    source_node: null
    target_node: "Gibe-III"
    kind: "x"
    flow_rate: # m^3/s
      - 53.61 # January
      - 47.81 # February
      - 40.70 # March
      - 58.80 # April
      - 88.19 # May
      - 191.98 # June
      - 582.05 # July
      - 942.25 # August
      - 592.34 # September
      - 316.45 # October
      - 138.75 # November
      - 69.97 # December
  - name: "Flow B"
    source_node: null
    target_node: "Koysha"
    kind: "x"
    flow_rate: # m^3/s
      - 21.44 # January
      - 19.12 # February
      - 16.28 # March
      - 23.52 # April
      - 35.28 # May
      - 76.79 # June
      - 232.82 # July
```

```
      - 376.90 # August
      - 236.94 # September
      - 126.58 # October
      - 55.50 # November
      - 27.99 # December
  - name: "Flow C"
    source_node: null
    target_node: "Omorate"
    kind: "x"
    flow_rate: # m^3/s
      - 15.01 # January
      - 13.39 # February
      - 11.40 # March
      - 16.46 # April
      - 24.69 # May
      - 53.75 # June
      - 162.97 # July
      - 263.83 # August
      - 165.86 # September
      - 88.61 # October
      - 38.85 # November
      - 19.59 # December
  - name: "Gibe-III Release"
    source_node: "Gibe-III"
    target_node: "Koysha"
    kind: "l"
  - name: "Koysha Release"
    source_node: "Koysha"
    target_node: "Kuraz Headworks"
    kind: "l"
  - name: "Canals to Kuraz Sugar Plantations"
    source_node: "Kuraz Headworks"
    target_node: null
    kind: "l"
    demand_flow: True
    demand_rate: # m^3/s
      - 98.77 # January
      - 56.58 # February
      - 15.43 # March
      - 0.00 # April
      - 17.49 # May
      - 17.49 # June
      - 29.84 # July
      - 54.53 # August
      - 48.35 # September
      - 75.62 # October
      - 105.45 # November
      - 104.42 # December
  - name: "Gibe-III Evaporation"
    source_node: "Gibe-III"
    target_node: null
    kind: "r"
    evaporation_flow: True
  - name: "Koysha Evaporation"
    source_node: "Koysha"
```

```yaml
    target_node: null
    kind: "r"
    evaporation_flow: True
  - name: "Main Channel"
    source_node: "Kuraz Headworks"
    target_node: "Omorate"
    kind: "r"
  - name: "Outflow to Lake Turkana"
    source_node: "Omorate"
    target_node: null
    kind: "r"
    demand_flow: True
    demand_rate:
      - 29.72 # January
      - 26.50 # February
      - 22.56 # March
      - 32.59 # April
      - 48.89 # May
      - 106.43 # June
      - 322.68 # July
      - 522.68 # August
      - 328.39 # September
      - 175.44 # October
      - 76.92 # November
      - 38.79 # December

objectives:
  - name: "Hydropower"
    kind: "Power Generation Maximization"
    target_node:
      - "Gibe-III"
      - "Koysha"
  - name: "Irrigation"
    kind: "Monthly Demand Deficit Minimization"
    quantile: 0.9 # 0-1
    target_flow: "Canals to Kuraz Sugar Plantations"
  - name: "Environment"
    kind: "Monthly Demand Deficit Minimization"
    quantile: 0.9 # 0-1
    target_flow: "Outflow to Lake Turkana"
```

D.2. Bathymetry Data

The bathymetry data for the Gibe III and Koysha reservoirs is provided in CSV files, the names of which are referenced in the configuration file. Top and bottom 5 lines of theses files are shown below.

D.2.1. Raw CSV Data from `Bathymetry Gibe-III.csv`

```
1    Volume,Surface,Head
2    14000000000,200000000,220
3    13986000000,199863614.7,219.93
4    13972000000,199727185.9,219.8599522
5    13958000000,199590713.6,219.7898566

     ...

998  56000000,4635225.839,39.40257495
999  42000000,3809652.795,36.38222589
1000 28000000,2889503.165,32.70735032
1001 14000000,1801256.04,27.86222671
1002 0,0,0
```

D.2.2. Raw CSV Data from `Bathymetry Koysha.csv`

```
1    Volume,Surface,Head
2    6000000000,119000000,178.5
3    5994000000,118918850.7,178.4432045
4    5988000000,118837675.6,178.3863703
5    5982000000,118756474.6,178.3294973

     ...

998  24000000,2757959.374,31.96981649
999  18000000,2266743.413,29.5192151
1000 12000000,1719254.383,26.53755469
1001 6000000,1071747.344,22.60639758
1002 0,0,0
```

E

Computational Setup

E.1. Hardware Specifications

The simulations were performed on a DelftBlue node with the following specifications:

- **Processor:** Intel Xeon Gold 6448Y (32 Cores, 60M Cache, 2.10 GHz) x 2 for a total of 64 cores
- **Memory:** 1 GB per core
- **Programming Language:** Python 3.10

E.2. Optimization Configuration Experiments

To select the best configuration in terms of population size, number of generations, and number of processes specified in the ProcessPool of Python's multiprocessing library, we conducted a series of experiments. Table E.1 shows the results of these experiments.

Table E.1: Comparative analysis of computational configurations for optimizing the Lower Omo-Gibe Basin model. This table presents the results of experiments conducted to determine the most efficient setup for the multi-objective evolutionary algorithm. Key performance indicators (KPIs) including Time per Function Evaluation (FE) per Core, Total Run Time, and CPU Utilization are compared across different combinations of Population Size, Number of Generations, and Number of Processes. These results inform the selection of the optimal computational configuration for subsequent optimization runs.

	Decision Variables			*KPIs*		
Run ID	**Population Size**	**Generations**	**Number of Processes**	**Time per FE per Core**	**Total Run Time**	**CPU Utilization**
A	128	625	128	60.6s	21.1h	94.90%
B	**128**	**625**	**64**	**58.4s**	**20.3h**	**96.15%**
C	640	125	128	61.2s	21.3h	92.83%
D	640	125	64	63.0s	21.9h	86.54%

We measured the Mean Time taken for completing one Function Evaluation (FE) per core and the CPU utilization to ensure efficient use of computational resources. A lower value of Time per FE and higher CPU utilization is desirable. Population size in these experiments was set as a multiple of the number of processes to ensure that each process gets an equal number of individuals to evaluate and the CPUs are utilized efficiently.

The results show that the best configuration is Run B, with a population size of 128, 625 generations, and 64 processes. Using this configuration, we achieved a CPU utilization of 96.15%. This configuration outperformed the other experiments for each KPI and was therefore selected for the subsequent optimization runs.

Bibliography

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. arXiv:1605.08695 [cs].

Amos, S., Mengistu, S., and Kleinschroth, F. (2021). Three decades of pastoralist settlement dynamics in the Ethiopian Omo Delta based on remote sensing data. *Human Ecology*, 49(5):525–537.

Avery, S. (2017). Fears over Ethiopian dam's costly impact on environment, people.

Avery, S. T. and Tebbs, E. J. (2018). Lake Turkana, major Omo River developments, associated hydrological cycle change and consequent lake physical and ecological change. *Journal of Great Lakes Research*, 44(6):1164–1182.

Biswas, A. (2004). Integrated Water Resources Management: A Reassessment. *Water International*, 29:8.

Blank, J. and Deb, K. (2020). Pymoo: Multi-Objective Optimization in Python. *IEEE Access*, 8:89497–89509.

Broomhead, D. and Lowe, D. (1988). Radial basis functions, multi-variable functional interpolation and adaptive networks. *ROYAL SIGNALS AND RADAR ESTABLISHMENT MALVERN (UNITED KINGDOM)*, RSRE-MEMO-4148.

Buhmann, M. D. (2003). *Radial Basis Functions: Theory and Implementations*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, Cambridge.

Carr, C. J. (2017). *River Basin Development and Human Rights in Eastern Africa — A Policy Crossroads*. Springer International Publishing, Cham.

Chaemiso, S. E., Abebe, A., and Pingale, S. M. (2016). Assessment of the impact of climate change on surface hydrological processes using SWAT: a case study of Omo-Gibe river basin, Ethiopia. *Modeling Earth Systems and Environment*, 2(4):1–15.

Cheng, R., Jin, Y., Olhofer, M., and Sendhoff, B. (2016). A Reference Vector Guided Evolutionary Algorithm for Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation*, 20.

Choi, Y. D., Roy, B., Nguyen, J., Ahmad, R., Maghami, I., Nassar, A., Li, Z., Castronova, A. M., Malik, T., Wang, S., and Goodall, J. L. (2023). Comparing containerization-based approaches for reproducible computational modeling of environmental systems. *Environmental Modelling and Software*, 167(105760).

Coello, C. C., Lamont, G. B., and Veldhuizen, D. A. v. (2007). *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer Science & Business Media. Google-Books-ID: 2murCij_wHcC.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. (2009). *Introduction to Algorithms*. The MIT Press, 3rd edition.

Dagne, H., Assefa, E., and Teferi, E. (2023). Mapping and Quantifying Land Degradation in the Omo-Gibe River Basin, South-Western Ethiopia. *African Geographical Review*, 0(0):1–15. Publisher: Routledge _eprint: https://doi.org/10.1080/19376812.2022.2164023.

Deb, K. and Jain, H. (2014). An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems With Box Constraints. *Evolutionary Computation, IEEE Transactions on*, 18:577–601.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197. Conference Name: IEEE Transactions on Evolutionary Computation.

Delft High Performance Computing Centre (DHPC) (2024). DelftBlue Supercomputer (Phase 2). https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase2.

Dondeyne, S., Kangi, G., Rosier, I., Kassa, H., and Van Orshoven, J. (2021). Climate, water and land resources : diversity, uses and changes. In *The Omo-Turkana Basin : cooperation for sustainable water management*, pages 11–36. Routledge.

Eiben, A. E. and Smith, J. E. (2003). *Introduction to Evolutionary Computing*. Natural Computing Series. Springer, Berlin, Heidelberg.

Foundation, A. S. (2015). What is Airflow? — Airflow Documentation.

Gebeyehu, B. M., Jabir, A. k., Tegegne, G., and Melesse, A. M. (2023). Reliability-weighted approach for streamflow prediction at ungauged catchments. *Journal of Hydrology*, 624:129935.

Gebre, S. L., Van Orshoven, J., and Cattrysse, D. (2023). Optimizing the Combined Allocation of Land and Water to Agriculture in the Omo-Gibe River Basin Considering the Water-Energy-Food-Nexus and Environmental Constraints. *Land*, 12(2):412. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

Gelalacha, A., Haile, M., and Tolera, A. (2024). Land suitability and crop water requirements for irrigated sugar cane in the Kuraz irrigation scheme, lower Omo basin, Ethiopia. *Irrigation and Drainage*.

Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., and Reed, P. M. (2016). Curses, Tradeoffs, and Scalable Management: Advancing Evolutionary Multiobjective Direct Policy Search to Improve Water Reservoir Operations. *Journal of Water Resources Planning and Management*, 142(2):04015050. Publisher: American Society of Civil Engineers.

Giuliani, M., Herman, J. D., Castelletti, A., and Reed, P. (2014). Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management. *Water Resources Research*, 50(4):3355–3377. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013WR014700.

Giuliani, M., Zaniolo, M., Block, P., and Castelletti, A. (2020). Data-driven control of water reservoirs using an emulator of the climate system. *IFAC-PapersOnLine*, 53(2):16531–16536.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., USA, 1st edition.

Hailu Woldegebrael, E. (2018). The Materialization of “Developmental State” in Ethiopia: Insights from the Gibe III Hydroelectric Development Project Regime, Omo Valley. *L’Espace Politique. Revue en ligne de géographie politique et de géopolitique*, (35). Number: 35 Publisher: Département de géographie de l’université de Reims Champagne-Ardenne.

Herman, J. and Giuliani, M. (2018). Policy tree optimization for threshold-based water resources management over multiple timescales. *Environmental Modelling & Software*, 99:39–51.

Hodbod, J., Stevenson, E. G. J., Akall, G., Akuja, T., Angelei, I., Bedasso, E. A., Buffavand, L., Derbyshire, S., Eulenberger, I., Gownaris, N., Kamski, B., Kurewa, A., Lokuruka, M., Mulugeta, M. F., Okenwa, D., Rodgers, C., and Tebbs, E. (2019). Social-ecological change in the Omo-Turkana basin: A synthesis of current developments. *Ambio*, 48(10):1099–1115.

Hutton, C., Wagener, T., Freer, J., Han, D., Duffy, C., and Arheimer, B. (2016). Most computational hydrology is not reproducible, so is it really science? *Water Resources Research*, 52:7548–7555. ADS Bibcode: 2016WRR....52.7548H.

International Rivers (2009). Ethiopia’s Gibe III Dam: Sowing Hunger and Conflict.

IPCC (2014). Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change. *IPCC*.

Jain, H. and Deb, K. (2014). An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point Based Nondominated Sorting Approach, Part II: Handling Constraints and Extending to an Adaptive Approach. *IEEE Transactions on Evolutionary Computation*, 18(4):602–622. Conference Name: IEEE Transactions on Evolutionary Computation.

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

Koutsoyiannis, D. and Economou, A. (2003). Evaluation of the parameterization-simulation-optimization approach for the control of reservoir systems. *Water Resources Research*, 39(6). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2003WR002148.

Lempert, R. J., Popper, S. W., and Bankes, S. C. (2003). Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis. Technical report, RAND Corporation.

Lukas, P., Melesse, A. M., and Kenea, T. T. (2023). Prediction of Future Land Use/Land Cover Changes Using a Coupled CA-ANN Model in the Upper Omo–Gibe River Basin, Ethiopia. *Remote Sensing*, 15(4):1148. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

Mathewos, Y., Abate, B., and Dadi, M. (2022). Performance of the CORDEX-Africa Regional Climate Model in capturing precipitation and air temperature conditions in the Omo Gibe River Basin, Ethiopia. preprint, In Review.

Merrick, B. (2018). The power of hydrology in the Omo-Gibe River Basin : Gibe III and flood retreat agriculture on the river omo. *Masters Thesis Series in Environmental Studies and Sustainability Science*.

Micotti, M. (2020). DAFNE Basemap for the Omo-Turkana Basin case study.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J. (2008). Climate change. Stationarity is dead: whither water management? *Science (New York, N.Y.)*, 319(5863):573–574.

Mota, J., Siqueira, P., Souza, L., and Vitor, A. (2012). Training radial basis function networks by genetic algorithms. volume 1.

Niggemann, F. and Graf, L. (2020). Turbidity Datasets for Lake Turkana based on Landsat.

Oki, T. and Kanae, S. (2006). Global hydrological cycles and world water resources. *Science (New York, N.Y.)*, 313(5790):1068–1072.

Pahl-Wostl, C. (2007). Transitions Towards Adaptive Management of Water Facing Climate and Global Change. *Water Resources Management*, 21:49–62.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.

Quinn, J. D., Reed, P. M., Giuliani, M., Castelletti, A., Oyler, J. W., and Nicholas, R. E. (2018). Exploring How Changing Monsoonal Dynamics and Human Pressures Challenge Multireservoir Management for Flood Protection, Hydropower Production, and Agricultural Water Supply. *Water Resources Research*, 54(7):4638–4662.

Quinn, J. D., Reed, P. M., and Keller, K. (2017). Direct policy search for robust multi-objective management of deeply uncertain socio-ecological tipping points. *Environmental Modelling & Software*, 92:125–141.

Rocklin, M. (2015). Dask: Parallel Computation with Blocked algorithms and Task Scheduling. *Proceedings of the 14th Python in Science Conference*, pages 126–132. Conference Name: Proceedings of the 14th Python in Science Conference.

Rosenstein, M. and Barto, A. (2001). Robot Weightlifting By Direct Policy Search. *IJCAI International Joint Conference on Artificial Intelligence*.

Schwenker, F., Kestler, H. A., and Palm, G. (2001). Three learning phases for radial-basis-function networks. *Neural Networks*, 14(4):439–458.

Studio Pietrangeli (2024a). Gibe III hydropower (Ethiopia).

Studio Pietrangeli (2024b). Koysha Hydropower (Ethiopia).

Voinov, A., Kolagani, N., McCall, M. K., Glynn, P. D., Kragt, M. E., Ostermann, F. O., Pierce, S. A., and Ramu, P. (2016). Modelling with stakeholders: next generation. *Environmental modelling & software*, 77:196–220. Publisher: Elsevier.

Vörösmarty, C. J., Green, P., Salisbury, J., and Lammers, R. B. (2000). Global Water Resources: Vulnerability from Climate Change and Population Growth. *Science*, 289(5477):284–288. Publisher: American Association for the Advancement of Science.

Webuild Group (2024a). Gibe III Hydroelectric Project | Webuild Group.

Webuild Group (2024b). Koysha Hydroelectric Project | Webuild Group.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3:160018.

Zaniolo, M., Giuliani, M., Bantider, A., and Castelletti, A. (2021a). Hydropower development: Economic and environmental benefits and risks. In *The Omo-Turkana Basin*. Routledge. Num Pages: 21.

Zaniolo, M., Giuliani, M., and Castelletti, A. (2021b). Policy Representation Learning for Multiobjective Reservoir Policy Design With Different Objective Dynamics. *Water Resources Research*, 57(12):e2020WR029329.

Zaniolo, M., Giuliani, M., Sinclair, S., Burlando, P., and Castelletti, A. (2021c). When timing matters—misdesigned dam filling impacts hydropower sustainability. *Nature Communications*, 12(1):3056. Number: 1 Publisher: Nature Publishing Group.

Zatarain Salazar, J., Kwakkel, J. H., and Witvliet, M. (2024). Evaluating the choice of radial basis functions in multiobjective optimal control applications. *Environmental Modelling & Software*, 171:105889.

Zatarain Salazar, J., Reed, P. M., Herman, J. D., Giuliani, M., and Castelletti, A. (2016). A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control. *Advances in Water Resources*, 92:172–185.

Zeff, H. B., Hamilton, A. L., Malek, K., Herman, J. D., Cohen, J. S., Medellin-Azuara, J., Reed, P. M., and Characklis, G. W. (2021). California's food-energy-water system: An open source simulation model of adaptive surface and groundwater management in the Central Valley. *Environmental Modelling & Software*, 141:105052.