

Article

Integrating Graph Data Models in Advanced Water Resource Management: A New Paradigm for Complex Hydraulic Systems

Jiabin Qiao ^{1,2} , Keyan Shen ¹ , Wenzhe Xiao ², Jiarun Tang ², Yurun Chen ² and Jing Xu ^{2,*}

¹ Hubei Key Laboratory of Intelligent Yangtze and Hydroelectric Science, China Yangtze Power Co., Ltd., Yichang 443000, China; qiaojiabin@hust.edu.cn (J.Q.); shen_keyan@ctg.com.cn (K.S.)

² School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China; xiaowz@hust.edu.cn (W.X.); tangjiarun@hust.edu.cn (J.T.); chen_yurun@hust.edu.cn (Y.C.)

* Correspondence: xujing@hust.edu.cn

Abstract: The efficient management of water resources is crucial due to increasing environmental and societal pressures, requiring advanced data handling capabilities. Traditional relational data models, while foundational, struggle to capture the complex relationships inherent in water systems, limiting analytical and decision-making effectiveness. To address these challenges, this paper proposes a comprehensive graph data model for water resource management. Our model includes conceptual, logical, and physical frameworks that enable the structured representation of hydraulic objects and their relationships. We categorize hydraulic objects, examine their spatial, structural, and operational interconnections, and construct an adaptable graph data model. This model provides the enhanced integration, mining, and analysis of hydraulic data, supporting intelligent, data-driven water management applications. Example applications demonstrate the model's utility in improving watershed management, hydropower scheduling, and uncovering latent relationships among hydraulic components through graph neural networks. This work establishes a robust data foundation for the sustainable and scientific management of water resources.



Academic Editor: Luís Filipe Sanches Fernandes

Received: 12 November 2024

Revised: 16 December 2024

Accepted: 17 December 2024

Published: 24 December 2024

Citation: Qiao, J.; Shen, K.; Xiao, W.; Tang, J.; Chen, Y.; Xu, J. Integrating Graph Data Models in Advanced Water Resource Management: A New Paradigm for Complex Hydraulic Systems. *Water* **2025**, *17*, 3. <https://doi.org/10.3390/w17010003>

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: water resource management; hydraulic systems; data modeling; graph data model

1. Introduction

Water resources are the foundation for human survival and development [1]. With global climate change, population growth, and economic development, the management demands for water resources are becoming increasingly complex and diverse [2]. Modern water resource management requires not only the rational allocation and effective utilization of water resources but also reliance on advanced data management technologies to achieve scientific decision-making and sustainable development. Hydraulic data models serve as core tools for information management, enabling the rapid acquisition, analysis, and utilization of data in complex water resource environments [3], thereby enhancing the scientificity and effectiveness of decision-making. The construction of efficient hydraulic data models is not only crucial for addressing the challenges of water resource management but also an important avenue for promoting the modernization of water resource management [4]. Thus, in-depth research into the construction and application of hydraulic data models is particularly urgent and significant.

Currently, the complexity and multidimensional characteristics of hydraulic data present numerous challenges for management and application [5]. For instance, the genera-

tion of diverse data sources such as water quality monitoring data, rainfall records, river flow rates, and groundwater levels not only results in vast amounts of data but also involves interactions across temporal, spatial, and multiple dimensions. According to statistics from the National Water Resources Monitoring Network, the sampling frequency at some water quality monitoring sites can reach several times daily, leading to exponential growth in data volume [6]. These data must be collected in real time and effectively integrated and analyzed. However, traditional relational databases often struggle to efficiently express the complex relationships between hydraulic data due to their fixed structure. For example, in water quality monitoring, the complexity of pollutant sources, watershed characteristics, and their interactions far exceeds the processing capabilities of traditional relational models [7]. This limitation hampers the ability of existing models to fully exploit the potential value of hydraulic data during data mining and analysis, consequently impacting the scientificity and efficiency of water resource management.

The introduction of graph data models offers new insights for addressing these issues. Graph data models, with their flexibility and expressive power, can effectively capture the diverse relationships among data and have achieved significant results in various fields, such as social network analysis, recommendation systems, and bioinformatics [8]. In recent years, some researchers have begun exploring the application of graph data models in the hydraulic field, for instance, analyzing flow networks and optimizing water resource allocation through graph models [9–11]. However, despite the preliminary demonstration of the potential of graph data models in water management, there remains a lack of systematic organization and analysis of hydraulic data, and a complete hydraulic graph data model has yet to be established. Existing research tends to focus on specific application problems, failing to comprehensively consider the diversity and complexity of hydraulic data, particularly with regard to their organization and structure. Therefore, in light of the complexity and multidimensional nature of hydraulic data, there is an urgent need to construct adaptive graph data models to better meet the diverse demands of modern water resource management.

To this end, this paper aims to introduce graph data models into hydraulic data management by systematically constructing a hydraulic graph data model in response to practical application needs. We design a hydraulic graph data model characterized by high expressiveness and flexibility based on the features of water resource management and explore typical applications of graph semantics on this foundation. Specifically, the main contributions of this paper include the following:

- We provide a systematic understanding of hydraulic objects' classification and relationships, which is crucial for building a robust data model.
- We develop a graph-based hydraulic data model that effectively accommodates the integration of diverse data types, facilitates efficient data mining, and supports complex data interactions.
- We propose several typical application examples of graph semantics based on this data model, aimed at promoting the intelligent and scientific development of water resource management.

Through this research, we hope to provide a theoretical foundation and practical guidance for the future development of hydraulic data management and to promote the sustainable management and utilization of water resources. The research presented in this paper not only offers new perspectives for the integration and analysis of hydraulic data but also contributes new ideas and methods to advance the modernization of water resource management.

This paper is organized as follows: Section 2 reviews the related work on data management for water resources, covering traditional and emerging data models. Section 3

provides a systematic analysis of hydraulic objects and their relationships, establishing a classification framework. In Section 4, we describe the construction of the hydraulic data model, detailing the conceptual, logical, and physical models. Section 5 discusses the discovery of implicit relationships among hydraulic objects, using topological and graph neural network methods to reveal hidden dependencies. Finally, Section 6 presents the conclusions and suggests directions for future research.

2. Related Work

2.1. Data Management for Water Resources

Effective water resource management relies heavily on robust data storage and management systems. To accommodate the unique characteristics of water-related data, various data models have been developed, each with distinct approaches to data modeling and representation, processing methods, and sharing capabilities. This subsection reviews key data models and highlights their features, advantages, and limitations.

2.1.1. Data Modeling and Representation

The ArcHydro data model [12], developed by the Environmental Systems Research Institute (ESRI), provides a comprehensive framework for managing hydrological and hydraulic data within the ArcGIS platform. By integrating spatial and temporal data, ArcHydro facilitates the representation of water systems, including river networks and watershed boundaries. Its compatibility with the Geographic Information System (GIS) enables sophisticated spatial querying and visualization, supporting standardized data sharing among researchers and water managers. This model has been instrumental in hydrological modeling and simulation, aiding data-driven decision-making in water management.

In contrast, HydroDB [13] is specifically designed for the management of hydrological and water quality data, specifically addressing the challenges posed by multi-dimensional attributes like time, location, and water quality indicators. HydroDB's flexibility enables the integration of real-time monitoring with historical data, making it particularly useful for regulatory compliance and environmental assessments. While ArcHydro excels in spatial data analysis, HydroDB's adaptable structure better accommodates diverse datasets, facilitating essential insights into water quality and compliance with governance standards.

Building upon these models, the watershed modeling framework (WMF) offers a suite of functionalities that are highly relevant to our discussion on data modeling and representation. The WMF [14] is a vital open-source tool for hydrological and geomorphological analysis in watershed management. It simulates water balance and quantifies erosion and sediment transport, offering detailed insights into water flow and land use impacts on watershed health. The WMF's integration with Quantum GIS (QGIS) enhances spatial data visualization and analysis, making it a valuable asset for understanding water system dynamics and supporting data models discussed in this paper.

2.1.2. Data Processing Methods

While ArcHydro and HydroDB are effective in organizing structured water resource data, they face limitations; as the volume, complexity, and heterogeneity of water-related data continue to grow, traditional relational databases may struggle to accommodate the diverse and dynamic nature of this information. This limitation has led to increased interest in Not Only Structured Query Language (NoSQL) databases (e.g., MongoDB, Cassandra), which provide more flexible storage solutions for unstructured and semi-structured data [15].

NoSQL databases provide significant advantages by supporting dynamic querying and real-time analytics, making them suitable for applications characterized by high data

variability. Their ability to accommodate diverse data types enables the integration of various data sources, enabling innovative approaches to water resource management. As organizations seek to harness the full potential of their data, transitioning to NoSQL systems can enhance the scalability issues present in traditional relational models. However, even NoSQL databases face challenges when handling complex interrelationships, such as those in large-scale hydrological systems.

2.1.3. Data Sharing Capabilities

Increasing demands for interoperability have led to the development of data-sharing frameworks like the WaterML [16] and Hydroinformatics data models [17], both of which are designed to facilitate seamless data exchange among diverse systems [18]. WaterML, an XML-based data exchange format, standardizes the representation of hydrological data, making it accessible across platforms and stakeholder groups. This common format enhances collaboration among researchers, policymakers, and water managers, ultimately improving data accessibility and usability.

Hydroinformatics data models, on the other hand, integrate traditional database frameworks with data warehousing techniques, allowing for the aggregation of diverse datasets from multiple sources. These models prioritize interoperability and support comprehensive water resource assessments, promoting a collaborative approach to water management across various sectors.

Existing models for water resource management provide essential tools for data storage, processing, and sharing, yet they fall short in managing the intricate, interdependent relationships characteristic of large-scale hydrological systems. Relational and NoSQL databases are not inherently suited to capture multi-dimensional dependencies, such as spatial and temporal interconnections among water bodies, hydraulic structures, and monitoring points. Consequently, there is a clear need for a data model that can accommodate these complex relationships and enable sophisticated analyses to support dynamic, data-driven decision-making. The transition from relational databases to more flexible systems like NoSQL has helped address some of these challenges, but capturing complex, multi-dimensional relationships between spatial, temporal, and environmental factors requires more advanced approaches.

2.2. Graph Data Models and the Transition to Advanced Frameworks

Graph data modeling has emerged as a powerful alternative to traditional data modeling, particularly in applications where relationships between data points are complex and dynamic [19]. Traditional relational databases organize data into structured tables, limiting their capacity to represent and explore intricate interdependencies. Graph data models, on the other hand, represent data as nodes (entities) and edges (relationships), which provides a flexible and intuitive way to model complex relationships. This structure is especially well-suited for applications focused on relational data, such as social networks, fraud detection, and recommendation systems, where graph data models can efficiently manage large volumes of interconnected data, support dynamic queries, and facilitate rapid relationship traversal. These attributes make graph databases effective tools for analyzing complex systems and supporting real-time analytics. As a result, graph data modeling has found widespread applications in various domains, including social networks [20], fraud detection [21], recommendation systems [22], and logistics optimization [23].

In the field of water resource management, graph data models offer innovative solutions for representing and analyzing the complex interrelationships among hydrological entities. By leveraging graph-based structures, researchers can gain insights into water distribution networks, pollutant migration pathways, and the interactions between water

bodies and infrastructure. For example, graph theory has been applied to model river networks, enabling the assessment of flow patterns and identification of critical nodes within the system [24]. Graph-based frameworks have also been used to optimize water resource allocation, where relationships between water sources, demand points, and distribution networks are represented, thereby facilitating more efficient resource management [25].

While existing studies highlight the benefits of graph data models in water resource management, there remains a need for a comprehensive framework to fully address the field's challenges. By establishing a more interconnected approach, graph data models can support the development of intelligent water resource management systems, ultimately providing the robust insights and adaptability needed for sustainable water management and effective policy-making.

3. Methods

3.1. Systematical Review on Hydraulic Objects and Their Interrelationships

Hydraulic objects refer to entities within water resource systems that possess functional and managerial significance. These entities include water resource infrastructure, natural water bodies, and other key elements associated with the scheduling and utilization of water resources. Hydraulic objects encompass various categories, and there often exist complex interactions and relationships among them. This complexity is manifested not only in direct physical connections but also in the multi-level impacts and feedback mechanisms within a system. A systematic analysis of the classification and characteristics of hydraulic objects, along with an in-depth exploration of their interrelationships, is a critical step for understanding and modeling hydraulic systems. It serves as a foundation for achieving precise scheduling, risk control, and resource optimization. Effectively identifying the nature of these objects and their complex relationships contributes to enhancing the overall efficiency of hydraulic systems, ensuring the scientific basis of decision-making and the sustainability of management.

3.1.1. Classification and Characterization of Hydraulic Objects

Hydraulic objects are core components of water resource management, encompassing a variety of natural entities and human-engineered facilities. Within the context of hydraulic engineering and water resource management, these objects exhibit significant diversity and complexity. Based on their physical characteristics, functional attributes, and roles in water resource management, hydraulic objects can be classified into three main categories: natural resource entities, hydraulic facilities, and non-facility elements, as shown in Figure 1. This classification framework establishes a foundation for the analysis and management of complex water resource systems, aiding in decision-making and effective resource management strategies.

Natural resource entities encompass fundamental units that are integral to the flow, distribution, and cycling of water within the natural environment, characterized by their inherent nature and dynamic variability. This category includes essential components such as watersheds, rivers, river segments, and cross-sections. Watersheds serve as crucial spatial units for water resource allocation and management, delineating areas where surface water converges, typically dictated by topographical features. They are foundational for water supply, collecting precipitation, surface runoff, and groundwater. Rivers, functioning as natural conduits, connect disparate watersheds and downstream water bodies while fulfilling vital roles in drainage, flood control, water supply, and ecological preservation. River segments represent specific divisions of rivers, defined by hydrological and geological attributes or management objectives, each exhibiting distinctive flow characteristics, water quality, and ecological functions. Cross-sections are employed for hydrological

monitoring and analysis, capturing instantaneous flow dynamics at designated locations. The interrelated dynamics of these natural resource entities are essential to effective water resource management, highlighting their significance as foundational elements within ecological systems.

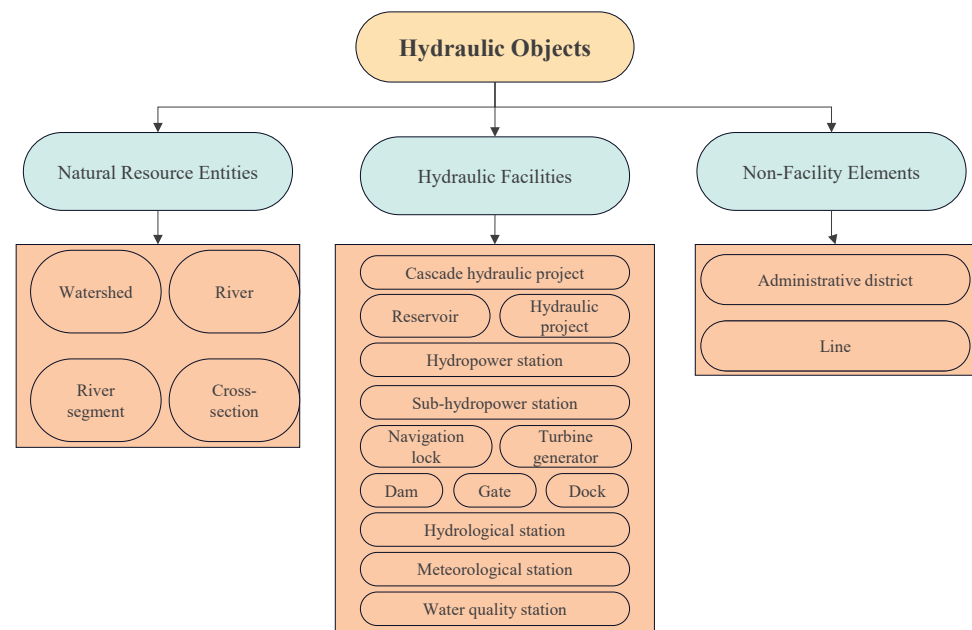


Figure 1. Classification framework for hydraulic objects.

Hydraulic facilities encompass all artificially constructed water facilities designed to control, store, and allocate water resources, forming a core component of modern water resource management systems. This category includes essential structures such as cascade hydraulic projects, reservoirs, hydropower stations, navigation locks, dams, hydrological stations, and water quality stations. Cascade hydraulic projects implement a series of hydraulic hubs along river corridors to achieve the segmented management and utilization of water resources, thereby simultaneously addressing multiple needs such as flood control, irrigation, and power generation. Reservoirs, as water storage facilities, provide regulated water supply during dry periods while also serving flood mitigation functions during wet seasons. Key indicators of reservoir effectiveness include storage capacity, engineering scale, and operational performance. Hydropower stations harness the potential energy of water to generate electricity, representing a significant source of clean energy. Their efficiency and sustainability are contingent upon the reliable supply of water resources. Navigation locks facilitate the safe passage of vessels at dams with significant water level differences, promoting waterway transport and logistics. Dams serve as critical infrastructure for flood control and water resource regulation, intercepting water flows to enable storage, electricity generation, and watershed management. Within the realm of hydraulic facilities, hydrological and water quality stations are vital for resource management. Hydrological stations conduct the long-term monitoring of key hydrological parameters, including flow rates, water levels, and precipitation, providing foundational data for flood forecasting, resource allocation, and water resource planning. The timeliness and accuracy of these data are crucial for effective water resource management and disaster risk reduction. Water quality stations focus on assessing the chemical and physical properties of water bodies, measuring parameters such as dissolved oxygen levels, pollutant concentrations, and temperature. By regularly monitoring water quality, these stations can promptly identify potential sources of contamination, ensuring the safety and sustainable use of

water resources. Through technical design and management operations, hydraulic facilities significantly enhance the efficiency of water resource utilization and the capacity for disaster prevention and mitigation.

Non-facility elements serve as the spatial and management framework for water resource management. Although lacking a physical form, these objects play a critical role in the allocation and decision-making processes related to water resources. This category primarily includes abstract entities such as administrative districts and transport lines, which facilitate the cross-regional distribution and effective coordination of water resources. Districts spatially delineate water resources based on geographical or management criteria, ensuring coordinated allocation and regulatory management within different jurisdictions. These divisions not only influence the delineation of management responsibilities but also significantly impact inter-regional water resource scheduling, ecological protection, and risk management. Transport lines are responsible for conveying electricity generated by hydropower stations, connecting these facilities to broader electrical grid systems and ensuring the seamless flow of energy across different regions. Through the clear definition of management boundaries and the optimization of resource transmission pathways, non-facility objects provide essential support for the effective operation of hydraulic engineering and the efficient distribution of resources.

3.1.2. Analysis of the Relationships Among Hydraulic Objects

The relationships among hydraulic objects are complex and diverse, forming an inter-dependent and influential system. These objects are not only physically connected but also functionally complementary, collectively ensuring the sustainable use of water resources and the balance of aquatic ecosystems. Analyzing the interrelationships among hydraulic objects involves examining three distinct aspects—spatial, functional, and operational ones—based on the data characteristics of the water industry, as shown in Table 1.

Table 1. Relationship between hydraulic objects.

Relationship Type	Specific Relationship
Spatial Relationship	Location Association
Structured Functional Relationship	Left/Right Bank Upstream/Downstream Main/Tributary Connection
Non-Structured Functional Relationship	Composition and Inclusion
Operational Relationship	Data Interaction Collaboration Retention Regulation Power Transmission

Spatial relationships primarily refer to the relative positions of hydraulic facilities in geographic and hydrological contexts and their mutual influences. These relationships encompass both the locations of facilities and their association, reflecting not only physical interdependence but also profoundly impacting hydrological behavior and management strategies. The locational relationship describes the physical embedding of facilities in specific geographic areas, determining their interactions with surrounding hydrological environments. For instance, the spatial arrangement of hydropower stations and reservoirs

directly affects water flow paths, velocities, and discharge rates. The association aspect illustrates the dependencies among hydraulic facilities that are spatially adjacent or in close proximity, typically interconnected within geographic or hydrological networks, thereby influencing each other's operations and scheduling. This is exemplified by the collaborative mechanisms established for water resource allocation among neighboring facilities.

Functional relationships emphasize the purposes and hierarchical dependencies between water facilities to support overall system functionality. Structured functional relationships, such as the left/right bank, upstream/downstream, and main/tributary connections, are essential in river networks. The left/right bank relationships illustrate the layout of hydraulic facilities along both sides of a river, a configuration closely linked to topographical and hydrological characteristics. This arrangement influences the distribution of functions and hydrodynamic behavior of the facilities; for example, the symmetrical design of hydropower stations on both banks can affect the direction and velocity of water flow. Upstream/downstream relationships enhance the interactive characteristics of longitudinal hydrological systems by regulating downstream hydrological conditions through the control of upstream flow. This dynamic plays a crucial role in flood control, pollution diffusion, and water resource utilization. Main/tributary relationships represent the linkages within a river system, where tributaries either supply or distribute water, forming a complex hydrological and resource distribution network. Non-structured functional relationships focus on compositional dependencies. For instance, hydraulic projects and their associated hydropower stations achieve synergies in functions such as flood control, electricity generation, and irrigation through the integration of different functional modules, thereby ensuring the stability and resilience of the overall system.

Operational relationships encompass the collaboration and interaction of hydraulic facilities during daily operations and management processes. Data interaction is primarily reflected in the real-time sharing and integrated management of hydrological data, meteorological information, and operational data, enhancing the efficiency of cooperation among facilities through the flow of information. This data interaction is particularly crucial in water resource allocation and risk warning, especially in flood prevention and the optimization of resource allocation, where data sharing between facilities aids in more precise decision-making and resource scheduling. Collaborative relationships further describe how hydraulic facilities work together to execute tasks during operations, achieving the optimal utilization of water resources within a region through coordinated efforts. The retention relationship pertains to facilities, particularly reservoirs and dams, that control the supply and use of water resources by intercepting and storing natural water flow. The essence of this relationship lies in how effectively these facilities can store and release water resources, ensuring a balance of water availability across different seasons and climatic conditions. Regulatory relationships refer to the dynamic control of water volumes by hydraulic facilities, ensuring the rational utilization and distribution of water resources. This involves balancing various demands through the regulation of water quantities. The power transmission relationship focuses on the transfer and scheduling of electricity between power stations and the grid, where flexible adjustments in power generation are made to meet the electricity needs of different regions and time periods, ensuring a grid load balance and the rational allocation of electrical resources.

Overall, spatial, functional, and operational relationships collectively form a complex and dynamic network for water resource management and regulation. A comprehensive understanding of these interconnected relationships improves planning, design, and sustainable water resource management practices.

3.2. The Conceptual Model

To develop a robust data model for water resource management, it is essential first to understand the core water resource entities and their interdependencies. Following a systematic review of these entities with a particular emphasis on hydraulic infrastructure, we now turn to construct the data model for advanced water resource management. This model will encapsulate the complex interactions among natural water resources, engineered facilities, and operational outputs, addressing the critical dimensions of resource allocation, real-time management, and sustainable utilization in water resource systems.

A conceptual model, at its core, provides an abstraction of a complex system by capturing essential entities, relationships, and functionalities within a structured framework. The goal of this model is to establish a flexible and scalable architecture that supports advanced data integration, analysis, and decision support within water resource management. By representing entities like hydraulic facilities and natural water resources along with their hierarchical and functional relationships, this model is designed to enhance operational efficiency, inform real-time decision-making, and promote sustainable management across administrative regions.

3.2.1. Layered Architecture of Conceptual Model

Rooted in hierarchical architecture, our data model employs a “three-tier framework” that decomposes key water resource entities into distinct but interconnected layers, each representing unique dimensions of the water management system, as shown in Figure 2. The Natural Resource Layer serves as the foundation, incorporating a spectrum of hydrological resources such as rivers (mainstream and tributaries) and other water bodies. These elements form the ecological basis upon which the infrastructure is constructed. The Hydraulic Facility Layer, central to this architecture, encompasses engineered assets such as hydropower stations, reservoirs, and ancillary structures like ship locks. Serving as the operational core, this layer enables the regulated control, distribution, and utilization of water resources. The Service Delivery Layer represents the end-point of infrastructure operations, where the generated resources (e.g., electricity) are distributed to administrative regions, aligning infrastructure output with regional service demands. This layer completes the model by connecting infrastructure function with end-user requirements.

3.2.2. Intra-Layer Relationships: Hierarchical Structuring and Functional Cohesion

Each layer exhibits its own internal hierarchy and functional organization, crucial for achieving efficient resource management. In the Natural Resource Layer, water bodies are subdivided into functional segments, such as river sections and cross-sections, allowing granular management based on hydrological parameters and geographic features. This segmentation provides an organized basis for dynamic monitoring, modeling, and localized intervention within the water system. Within the Hydraulic Facility Layer, hydropower stations and their associated components are structured as interconnected hubs, functioning as operational nodes within a networked control system. These hubs are integrated into cascading frameworks, facilitating the multi-site coordination and optimization of water flow, storage, and energy generation. This intra-layer structuring ensures functional cohesion, where each object within a layer contributes strategically to the layer’s overarching role.

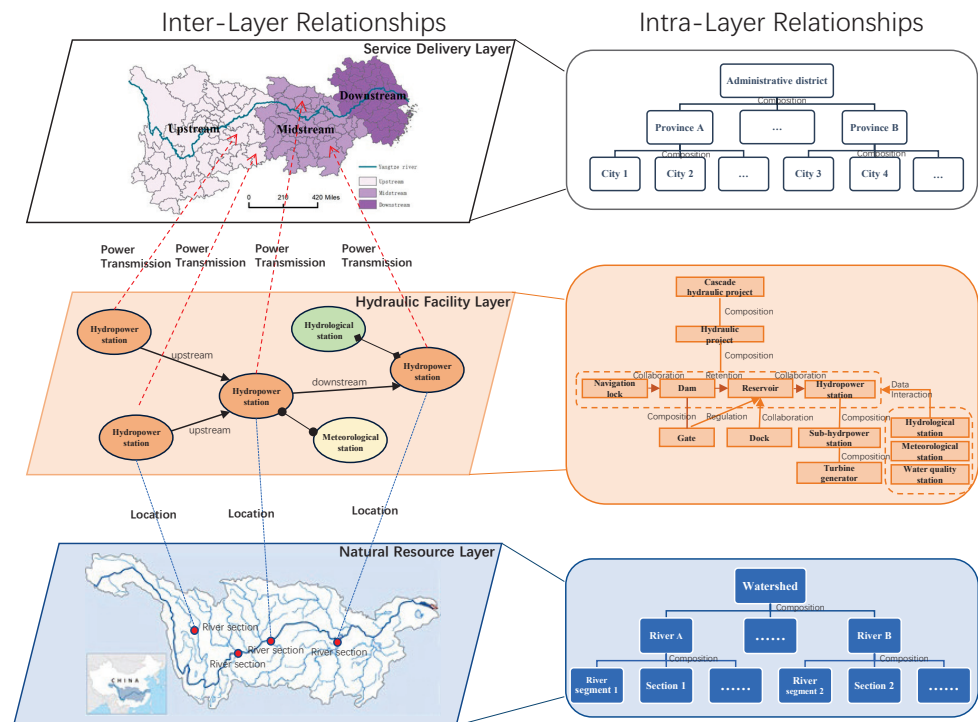


Figure 2. Layer architecture of conceptual model.

3.2.3. Inter-Layer Relationships: Spatial Synchronization and Operational Continuity

The model's inter-layer relationships establish a dynamic interface between natural resources, infrastructure, and administrative outputs. The geographical location of the hydraulic facilities is linked to the specific section of the river, and a direct spatial synchronization link is established between the Natural Resource Layer and the Hydraulic Facility Layer. This spatial alignment facilitates coordinated flow regulation, resource distribution, and environmental stewardship. Further, the operational outputs of the Hydraulic Facility Layer extend to designated administrative zones within the Service Delivery Layer, ensuring that resource distribution (e.g., electricity) meets specific regional demands. These inter-layer connections leverage spatial and functional continuity, creating a networked data model where natural resources, infrastructure, and operational output coalesce to support real-time, data-driven resource management.

This three-tier conceptual model offers a sophisticated, multi-dimensional approach to data representation and management in the field of water resources. By encapsulating natural resources, structured facilities, and service delivery layers, the model provides a flexible yet detailed framework for managing complex, interdependent water resource systems. This architecture, rooted in spatial synchronization and functional hierarchies, empowers decision-makers with a high-resolution view of resource dynamics and operational interdependencies. Such a model paves the way for innovative, responsive, and sustainable water resource management solutions.

3.3. The Logical Model

A graph data structure serves as a mathematical framework for representing objects and their interrelationships, commonly utilized for modeling and analyzing complex networks. Composed of nodes and edges, a graph illustrates entities as nodes and the relationships between them as edges. Each node and edge is further characterized by specific properties, facilitating a detailed understanding of entities and their interactions. Compared to relational data models, graph data structures offer significant advantages

in managing intricate relationships. The inherent flexibility of graphs enables a natural representation of multi-dimensional relationships and enhances query efficiency within complex networks. When addressing large-scale associative data, graph structures provide improved expressiveness and performance. Accordingly, we aim to construct a logical model using a graph data structure.

3.3.1. Attribute Description of Hydraulic Objects

In a hydraulic system, attributes refer to the information used to describe the state or characteristics of specific hydraulic objects. These attributes are typically categorized into two main types: basic data attributes and time series data attributes. Basic data refers to relatively fixed descriptive information that does not change frequently and is used to define and describe the fundamental characteristics of hydraulic objects over the long term. This type of data encompasses several aspects, including the following:

1. Identification information: This refers to characteristics that ensure the uniqueness and clarity of each hydraulic object within the system, including unique identifiers (IDs) and names.
2. Characteristic information: This describes the inherent attributes, physical properties, or functions of a hydraulic object, such as total reservoir capacity and regulation capacity.
3. Temporal information: This captures key time points associated with hydraulic objects, such as effective dates and expiration dates.

On the other hand, time series data attributes refer to parameters that change dynamically over time, such as water volume and electricity generation. These data typically represent the state of or variations in hydraulic objects at different time points, reflecting the trends of specific phenomena or parameters along the temporal dimension. The detailed definition of the attributes, A , of hydraulic objects is as follows:

$$A = \{B_i, B_f, B_t, S(t)\}, \quad (1)$$

where B_i is identification information, B_f is characteristic information, B_t is temporal information, and $S(t)$ is time series data.

In a hydraulic system, the attributes associated with different hydraulic objects exhibit variability. While identification information and temporal information often show a degree of commonality across most hydraulic objects, identification data generally comprise five essential attributes: a unique identifier (ID), name, spatial dimension, longitude, and latitude. These elements ensure both the uniqueness and geographic accuracy of hydraulic objects within the system. Temporal information includes two fundamental attributes—the effective date and expiration date—serving to document the lifecycle of hydraulic objects and reflecting their valid status during specific time intervals. Conversely, the characteristic information and time series data for different hydraulic objects can vary considerably. For instance, the characteristic information for a cascade hydraulic project encompasses total reservoir capacity, regulating capacity, and ownership details, while its time series data track annual cumulative electricity generation, total power output, transmission power, and total active power supplied. In contrast, the characteristic information for a cross-section primarily includes the section type and distance from the dam, with time series data reflecting the water level, flow rate, sediment concentration, and water temperature.

3.3.2. Analysis of Logical Structure

In the hydraulic data model, hydraulic objects are typically represented as nodes corresponding to the respective entities. Each node consists of labels and properties,

with node labels identifying the types of hydraulic objects, such as hydropower stations, reservoirs, and dams. Node labels facilitate the rapid indexing of similar types of hydraulic objects. The properties of nodes store the attributes corresponding to each entity. The identification information and temporal information attribute types of different hydraulic objects exhibit commonality, specifically in the five basic identification attributes and two fundamental temporal attributes. However, the characteristic information and time series data attribute types can vary significantly.

For instance, hydropower station nodes, labeled as “HPS”, include characteristic information attributes such as the total installed capacity, number of turbines, and maximum unit power, which describe the facility’s capacity and operational structure. Their time series data attributes include electricity generation, electricity loss, water utilization efficiency, and pressure differentials, reflecting time-varying operational metrics. Cross-section nodes, labeled as “Cross-Section”, represent hydrological observation points and characteristic information attributes like the type and distance from the dam, alongside time series data attributes such as the flow rate, water level, temperature, and sediment concentration. These properties provide detailed insights into river behavior. Administrative district nodes, labeled as “Administrative District”, represent geopolitical regions and contain characteristic information attributes like the area and administrative code, while their time series data attributes, such as GDP, reflect economic activity over time.

The relationships between entities are modeled through edges. Edges also consist of labels and properties. Edge labels indicate the type of relationship between entities, such as inclusion, collaboration, and data interaction. Unlike the properties of nodes, the properties of edges store additional information regarding the relationships themselves. This additional information includes an identifier (ID) for uniquely identifying specific edges and a property indicating the direction of the relationship, whether it is unidirectional or bidirectional. The edge structure also stores two node references, the starting node reference and the ending node reference, which point to the two nodes connected by the edge, ensuring a complete description of the associations between entities. Each reference is identified by the unique identification information B_i of the nodes, ensuring the accuracy and uniqueness of the edge connections.

For example, the edge labeled “Locate_at” represents the spatial relationship between a hydropower station and a specific cross-section of the river. In this case, the starting node might be a hydropower station (e.g., HPS_001), and the ending node would be a cross-section node (e.g., CS_001) in a specific location along the river. The “Locate_at” edge would include properties such as the distance from the hydropower station to the cross-section, reflecting the physical proximity and spatial relationship between these entities. Another important edge type is “Collaborate”, which denotes a functional relationship between two nodes, such as collaboration between a hydropower station and a reservoir. In this case, the starting node could be a hydropower station (e.g., HPS_001), and the ending node would be a reservoir (e.g., Reservoir_001). At the same time, there is a corresponding reverse edge, and the two together form a two-way relationship in the edge. The Collaborate edge might include properties like operational coordination or resource sharing, indicating how the two entities cooperate to optimize water resource utilization and electricity generation.

The detailed definitions of the node structure and edge structure within the graph are as follows:

$$Node = \{L, P_n\}, \quad (2)$$

$$Edge = \{L, P_e, N_{from}, N_{to}\}, \quad (3)$$

where L is the label, P_n is the node property, P_e is the edge property, N_{from} is the reference to the starting node of the edge, and N_{to} is the reference to the ending node of the edge.

3.3.3. Construction of Logical Model

Based on the analysis of the conceptual model and the graph logical structure, this paper constructs a graph logical model oriented toward water resource management, as illustrated in Figure 3. In this model, various types of relationships can be established among hydraulic nodes. These relationships can be quickly differentiated and identified based on labels, creating a clear and structured relationship network.

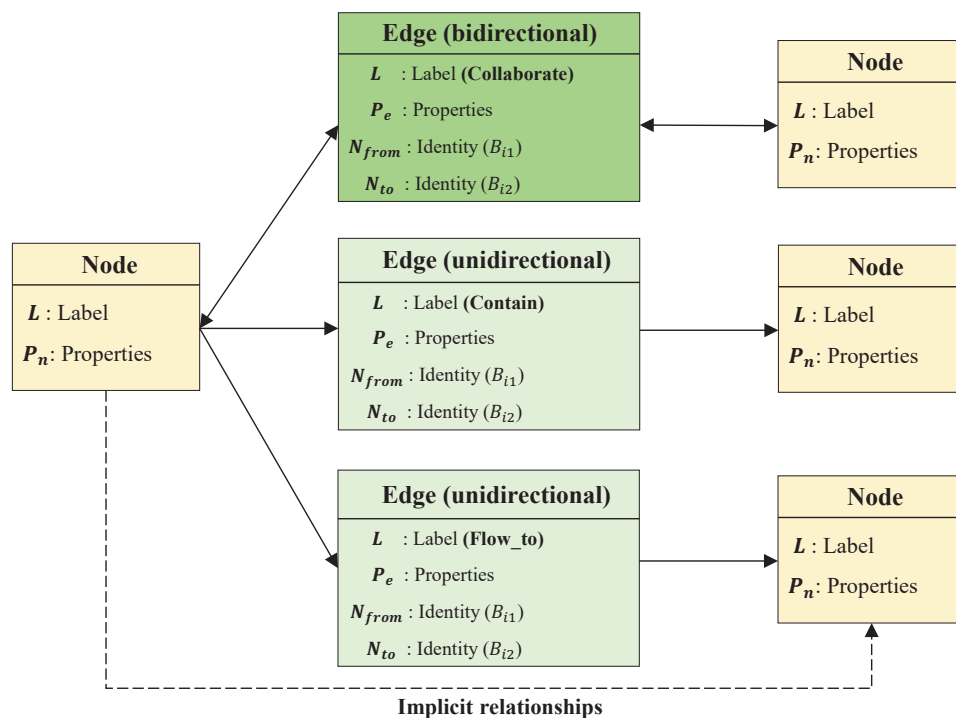


Figure 3. Graph logical model for water resource management.

The model offers high operational flexibility, supporting the convenient addition and deletion of nodes and edges. By manipulating node references, individual hydraulic objects can be easily added or removed, or relationships between objects can be updated without affecting the overall network structure. This capability allows the model to dynamically adjust in response to real-world water resource management needs, effectively enhancing management flexibility and responsiveness.

The model supports both unidirectional and bidirectional relationship representations, allowing for flexible adaptation to various management needs. Unidirectional edges indicate a directed dependency or association between the starting node and the ending node, such as the relationship where a hydropower station contains its sub-hydropower stations. Bidirectional edges are used to represent mutual associations between nodes, such as the collaboration relationship between a hydropower station and a reservoir, indicating a reciprocal benefit in resource sharing or task execution. Bidirectional relationships are formed by combining two unidirectional relationships, where the starting node reference N_{from} and the ending node reference N_{to} are symmetrically opposite. The Direction property of both edges is set to “bidirectional”, collectively representing the bidirectional connection between the nodes.

The logical model established in this paper has two special edge objects, namely the river segment edge and hidden edge. For river segment objects, the primary focus of research is often on flow direction relationships, which leads to their abstraction as edge structures within the graph. The endpoints of these edges connect the upstream and downstream cross-sections of the river segment, with the edge directed from the upstream

cross-section to the downstream cross-section. The edge label signifies the flow direction of the river segment, and the direction is set to “unidirectional”, with N_{from} representing the ID of the upstream cross-section and N_{to} representing the ID of the downstream cross-section. This abstraction of river segments as edge structures not only enhances the intuitiveness of the data representation within the river segment network but also improves the efficiency of node queries based on these connections. Such an approach significantly enhances the accuracy of flow direction analysis and query performance in complex hydraulic networks, particularly demonstrating notable advantages in the processing and visualization of large-scale river segment network data.

The implicit relationships within the graph are also revealed through the edges between nodes. For instance, different edge labels can identify potential collaborative or competitive relationships between nodes, which is crucial for understanding the complex flows of resources and information exchange within hydraulic systems. The direction property of the edges clarifies the flow of relationships, aiding in the identification of the pathways for resource or information transfer.

3.4. The Physical Model

The physical model is the concrete implementation of the logical model, which converts the entity relationships in the logical model into the actual database structure, and by building the physical model, the abstract concepts and business requirements in the logical model can be converted into the actual storage structure that can be executed in the database. Based on the graph data conceptual model and logical model oriented toward water resource management, this paper constructs a graph data storage system for hydraulic objects using the Neo4j graph database mechanism. A brief example of the hydraulic objects stored in the Neo4j database is shown in Figure 4. In Figure 4, HPS stands for hydropower station, TG stands for turbine generator, HLS stands for hydrological station, WQS stands for water quality station, and MS stands for meteorological station.



Figure 4. Graph physical model for water resource management.

The storage design of the database categorizes different entities into various node types based on the types of hydraulic objects, connecting these nodes through edges to represent the relationships among hydraulic objects. Each node and edge includes corresponding labels and properties, ensuring that both basic data and time series data for the hydraulic objects are fully preserved in physical storage. This model facilitates the

efficient querying and manipulation of hydraulic objects, providing a comprehensive and visual data foundation for the management and maintenance of hydraulic projects.

In the Neo4j database, both node and edge labels are stored as String. The specific properties and their storage formats are outlined in Tables 2 and 3. Due to significant differences in the characteristic and timing information of various entities, node features and timing data are stored using Map, where the keys represent different types of properties. The direction of edges is stored as a Boolean value, which determines whether the edge is unidirectional or bidirectional.

Table 2. Table of node properties.

Property Name	Type	Description
ID	String	Unique identifier
Name	String	Entity name
Latitude	Float	Latitude coordinate
Longitude	Float	Longitude coordinate
Feature	Map	Characteristic information
Timing	Map	Timing information
StartDate	DateTime	Effective time
EndDate	DateTime	Expiration time

Table 3. Table of edge properties.

Property Name	Type	Description
ID	String	Unique identifier
FromNodeID	String	A reference to the start node
ToNodeID	String	A reference to the end node
Direction	Boolean	Unidirectional or bidirectional

In the field of hydraulic engineering, the construction of a physical model provides a foundational basis for subsequent forecasting, simulation, and decision-making processes. This physical model enables more efficient processing and analysis of large-scale hydraulic facility data, particularly during complex water resource scheduling and management activities. Additionally, the model supports multidimensional querying and analysis, assisting managers in better understanding the interactions among various facilities. This insight can help optimize scheduling strategies and enhance water resource utilization efficiency, laying the groundwork for future intelligent water management systems.

4. Results

The graph data model proposed in this paper demonstrates strong applicability in various aspects of water resource management. Its versatility extends to multiple applications, including the efficient storage and management of hydraulic data, the discovery of hidden relationships between hydraulic objects, and the mining of data relationships using graph neural networks (GNNs).

4.1. Storage and Management of Water Conservancy Data

A water resource system typically encompasses data from multiple dimensions, such as water volume, water quality, climate, watersheds, and infrastructure (e.g., hydropower stations, dams, pumping stations), and data from various administrative regions. These data are not only vast in quantity but also highly interconnected and dynamically changing.

Traditional relational database models often struggle to effectively handle these complex, multidimensional relationships. Efficiently storing and managing these data, while ensuring their integrity, accuracy, and timeliness, is of critical importance for effective water resource management.

The storage and management of hydraulic data, as proposed by the graph data model in this paper, provides a flexible and efficient method for handling the complex relationships inherent in water resource systems. A specific application of water conservancy data integration in the Yangtze River Basin based on the graph data model is shown in Figure 5; the nodes and edges of different colors in the figure represent different types, and the text in the nodes is the real name of the node entity.

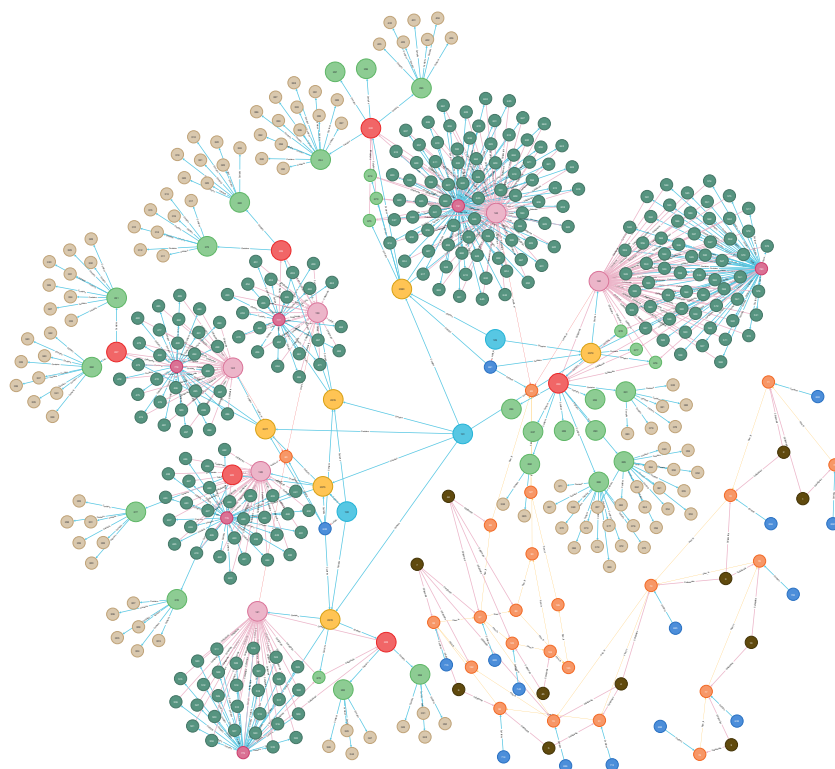


Figure 5. Water conservancy data integration in the Yangtze River Basin based on the graph data model.

After obtaining data related to the hydraulic entities from the Natural Resource Layer, Hydraulic Facility Layer, and Non-facility Layer, the graph-based approach allows for the effective storage and easy management of these entities and their interactions. First, each hydraulic entity is represented as a node in the graph, with the node label corresponding to the entity type (e.g., hydropower station, reservoir) and the node properties storing relevant attributes specific to each entity. For example, a hydropower station node might contain attributes such as the installed capacity, number of turbines, and location. Next, the relationships between specific hydraulic entities are modeled using edges. For instance, if hydropower station 1 is located in cross-section 1, an edge is established between these two nodes. The edge label is set as “Located_at”, and the edge properties include relevant information about the spatial relationship, such as the distance between the station and the cross-section.

Finally, the graph data model provides a comprehensive storage structure that integrates all hydraulic data, enabling the effective management of the entire system. In Figure 5, 467 entity nodes are constructed, including 6 hydropower station nodes, 254 gate nodes, 11 section nodes, and other nodes, as well as 798 direct connection edges be-

tween nodes, including 6 location relationships between hydropower stations and sections and 321 data cooperation relationships among hydropower stations, reservoirs, docks, etc. Tables 4 and 5 show the properties of the nodes and edges, in which the properties of the hydropower station node and the “Flow_to” edge are given, and the attribute information of the water conservancy entity is stored reasonably and regularly. Finally, the data integration of the water conservancy data is realized. By using nodes and edges to represent both entities and their interconnections, the graph structure ensures that the data are easily accessible, flexible for updates, and scalable for future expansions, thereby facilitating the efficient management of complex hydraulic systems.

Table 4. Table of hydropower station node properties.

Property Name	Value
ID	HPS001
Name	Three Gorges Hydropower Station on the Yangtze River
Latitude	30.821165
Longitude	111.00882
Feature_installedCapacity	22.5
Feature_numberOfTurbins	32
Timing_electricityGeneration	Eg (t)
Timing_electricityLoss	El (t)
Timing_pressureDifferential	Pd (t)
Timing_waterUtilizationEfficiency	Wue (t)
StartDate	2003-06-01T00:00:00Z
EndDate	null

Table 5. Table of “Flow_to” edge properties.

Property Name	Value
ID	Flow_to001
FromNodeID	Section002
ToNodeID	Section005
Direction	unidirectional
Length	43.5 km

4.2. Discovery of Hydropower Stations’ Topological Networks Based on Upstream and Downstream Relationships

In practical water resource management applications, the upstream and downstream relationships of hydropower stations play a crucial role in watershed scheduling and power generation planning. For instance, in the operation of a cascade hydropower station group, the power generation schedule of a hydropower station is not only influenced by its own reservoir storage levels but also depends on the operational status and discharge plans of upstream stations. If an upstream station suddenly increases its water release, the downstream station must promptly adjust its power generation and drainage to prevent dam failure or power loss. The timely and accurate identification of the upstream and downstream topological relationships among hydropower stations is essential for enhancing the efficiency of water resource utilization and ensuring the safety of power scheduling.

The upstream and downstream topological relationships of hydropower stations are established based on their connections with cross-sections. Each hydropower station is located at a specific cross-section, and the upstream–downstream relationships among cross-

sections are determined by the natural flow direction of river segments. This implies that the geographical and directional relationships between hydropower stations, cross-sections, and river segments dictate the upstream and downstream connectivity among the stations. By analyzing the connection between cross-sections where hydropower stations are situated and river segments, it is possible to derive indirect upstream and downstream relationships among hydropower stations. This provides essential data support for comprehensive scheduling within a watershed. The pseudocode for the algorithm is shown in Algorithm 1.

Algorithm 1 Establishment of upstream and downstream relationships

Require: Set of Hydropower Stations H , Set of River Sections R

Ensure: Upstream and downstream relationships among hydropower stations

```

1: for each hydropower station  $h_1 \in H$  do
2:   Let  $r_1 \leftarrow$  the section where  $h_1$  is located
3:   for each section  $r_2 \in$  find downstream sections of  $r_1$  do
4:     Let  $h_2 \leftarrow$  the hydropower station located at  $r_2$ 
5:     if  $h_2$  exists and the relationship  $(h_1, h_2)$  is not yet established then
6:       Establish the upstream-downstream relationship between  $h_1$  and  $h_2$ 
7:       Add  $(h_1, h_2)$  to the set of established relationships
8:     end if
9:   end for
10: end for

```

The algorithm constructs upstream and downstream relationships by traversing each hydropower station system. For each hydropower station h_1 , it identifies the corresponding cross-section r_1 based on its location relationship with that cross-section. Subsequently, it recognizes the downstream cross-section r_2 according to the flow direction relationship between cross-sections and searches for any potential downstream hydropower stations h_2 located in these downstream cross-sections. If a downstream hydropower station exists and the upstream–downstream relationship has not yet been established, the algorithm creates the relationship between the two hydropower stations and adds it to the set of established relationships. By thoroughly traversing all hydropower stations and cross-sections, this algorithm accurately constructs and records the upstream and downstream relationships among hydropower stations, providing a reliable data foundation and decision support for the efficient management of watershed water resources and the optimization of hydropower station scheduling.

The initial and resulting diagrams of the discovered upstream and downstream topological relationships for hydropower stations are shown in Figure 6. In Figure 6a, two hydropower stations (HPS_1, HPS_2) are located in two distinct cross-sections, which are interconnected through clear flow direction relationships, while there are no direct connections between the hydropower stations. The results of applying Algorithm 1 to establish the upstream and downstream relationships among the hydropower stations are shown in Figure 6b. An accurate representation of the upstream and downstream relationships is constructed, forming a topological structure among the hydropower stations. The clarification of these relationships allows each hydropower station to more precisely predict and respond to changes in the water flow, enabling more coordinated scheduling and control. The efficient transmission of data and collaborative processing between upstream and downstream hydropower stations facilitates timely optimization adjustments within the decision-making system, significantly enhancing the efficiency of data sharing and operational processing.

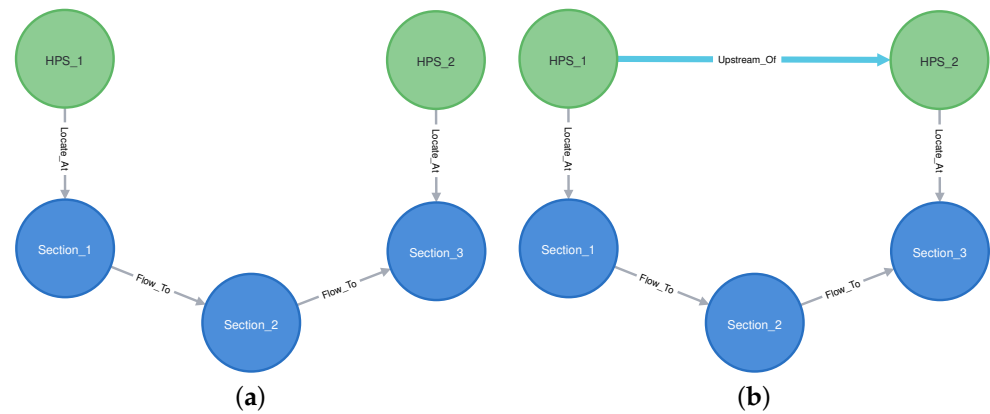


Figure 6. Discovery of upstream and downstream topological relationships for hydropower stations. (a) Diagram of relationships between hydropower stations and cross-sections. (b) Constructed upstream and downstream topological relationships of hydropower stations.

4.3. Relationship Mining of Hydraulic Objects Based on Graph Neural Networks

As the complexity of water resource management increases, traditional data mining techniques struggle to meet the demands for the in-depth analysis of relationships among hydraulic objects. Graph neural networks (GNNs) [26], an emerging deep learning approach, leverage their superiority in processing graph data to effectively uncover implicit relationships among hydraulic objects. GNNs facilitate information transmission and aggregation through node features and adjacency relationships, enabling the representation learning of nodes. This process is particularly well-suited for mining relationships among objects in the hydraulic domain, such as hydropower stations, rivers, and cross-sections.

In the hydraulic system, there exists a complex network of relationships among various objects. These relationships are not only reflected in the directly connected hydraulic objects but can also form deeper dependencies through indirect relationships. For example, the power generation capacity of a hydropower station may be influenced by upstream precipitation, while changes in downstream water levels can, in turn, affect the ecological balance of the river. By utilizing graph neural networks (GNNs), these intricate relationships can be modeled as a graph structure. During the process of information propagation and aggregation, GNNs effectively capture the relational features between nodes, allowing each node to learn not only its own characteristics but also to take into account the features of its neighboring nodes. With the stacking of multi-layer GNNs, nodes gain more comprehensive contextual information, thereby enhancing the overall performance of the model. The pseudocode for the algorithm is shown in Algorithm 2.

The algorithm employs graph neural networks (GNNs) to conduct in-depth mining of relationships among hydropower station objects. Initially, the algorithm initializes the weight matrix W of the GNN layers, which is used for the subsequent linear transformations of node features and information propagation. During each training epoch t , for each node v , the algorithm aggregates the features of neighboring nodes from the neighbor set $N(v)$. This aggregation is performed by applying the weight matrix $W^{(t)}$ for weighted summation, followed by the application of an activation function σ to extract nonlinear features. This process yields the feature representation $H_v^{(t)}$ for each node in the current epoch t , reflecting a comprehensive integration of the node's structural information and the features of its neighbors within the graph.

Next, the algorithm performs a reinforcement update on the node features by adding the new features $H^{(t)}$ to the features from the previous epoch. This process further enriches the node representations. At the end of each epoch, the loss function L is computed based on the actual relationships between nodes, serving to measure the discrepancy between

the current representations and the target representations. The weight matrix W is then adjusted through a backpropagation algorithm to optimize the model's representational capacity and robustness. Through multiple iterations, the algorithm continuously strengthens the expression of relationships between nodes, ultimately returning the feature representations H for all nodes, which encapsulate the potential associations among them. This rich embedded information lays the groundwork for subsequent relationship mining among hydropower station objects.

Algorithm 2 Graph neural network analysis for hydropower object relationship mining

Require: Graph G with nodes N and edges E .

Require: Node features X and adjacency matrix A .

Require: Hyperparameters: learning rate α , epochs T .

Ensure: Node representations H .

1: Initialize weights W for GNN layers.

2: **for** each epoch t in 1 to T **do**

3: **for** each node $v \in N$ **do**

4: Aggregate neighbors' features:

$$H_v^{(t)} \leftarrow \sigma \left(\sum_{u \in \mathcal{N}(v)} W^{(t)} H_u^{(t-1)} \right)$$

5: **end for**

6: Update node features:

$$H^{(t)} \leftarrow H^{(t)} + H^{(t)}$$

7: Compute loss L using ground truth relationships.

8: Update weights W by backpropagation.

9: **end for**

10: Return node representations H .

An application of Algorithm 2 is shown in Figure 7. In Figure 7a, hydropower stations supply electricity to administrative districts via transmission lines, with the power demand of each district collectively met by multiple hydropower stations. When an administrative district's power demand fluctuates, it is essential for the hydropower stations to respond rapidly by adjusting their power output to ensure accurate demand matching. However, if the distribution of power supply across hydropower stations is not properly managed during the scheduling process, this may result in inadequate supply, overload, or situations where a hydropower station's transmission demand exceeds its actual capacity, thereby compromising system stability.

To address these challenges, we employ a graph neural network (GNN) to analyze the hidden relationships between hydropower stations and administrative districts. Using Algorithm 2, the model determines the real-time allocation of electricity in response to fluctuations in the power demand of administrative districts, as shown in Figure 7b. The implicit power demand allocation relationships between administrative districts and hydropower stations enable the precise matching of the electricity demand across regions under dynamic conditions, while also preventing power wastage or excessive scheduling.

This method, based on the GNN model, effectively ensures that the electricity supply is distributed optimally and efficiently across the system. It significantly enhances the system's capability to respond to sudden fluctuations in demand, improving both the stability and reliability of the power distribution network.

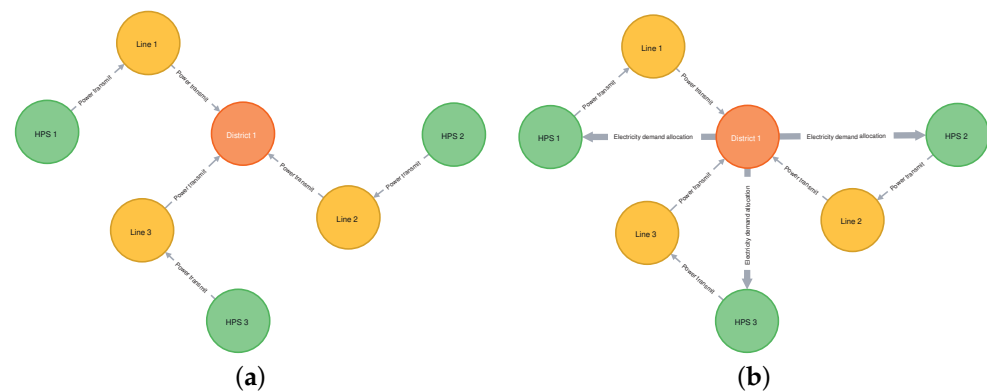


Figure 7. Implicit relationships found in electricity demand allocations for hydropower stations and administrative districts. (a) Diagram of relationships between hydropower stations and administrative districts. (b) Constructed electricity demand allocation relationships between hydropower stations and administrative districts.

5. Discussion

The complexity and interconnectivity of modern hydrological systems necessitate advanced data modeling approaches that extend beyond the capabilities of traditional relational models. These conventional systems often fail to effectively capture the intricate spatial, temporal, and operational interrelationships inherent in water resource systems. Addressing this limitation, the proposed graph-based data model leverages graph theory to provide a comprehensive, dynamic, and relationship-centric framework. This model enables the representation of hydraulic entities as nodes and their interactions as edges, inherently accommodating the multi-dimensional dependencies within water systems. By assigning labels and properties to nodes and edges, it facilitates detailed attribute management, allowing for the granular representation of data. Such capabilities are crucial for encapsulating the spatial, temporal, and operational nuances of hydrological systems while supporting advanced semantic analysis and query optimization, both essential for real-time decision-making and adaptive resource management.

The graph-based model serves as an innovative foundation for water resource data management, aligning seamlessly with established frameworks such as the watershed modeling framework (WMF) and semantic web technologies. Functioning as a foundational data organization layer, akin to database technologies, it provides robust support for these frameworks. Its ability to efficiently organize multi-dimensional relationships within localized water systems contributes significantly to broader harmonization efforts championed by the WMF. Moreover, the model's compatibility with semantic web technologies enhances data representation by leveraging shared ontologies, accelerating semantic querying and reasoning. This integration facilitates higher-level interoperability and advanced analytics, bridging the gap between the organizational strengths of the graph model and the interpretative capabilities of existing frameworks.

The scalability and versatility of the proposed model position it as a viable solution for large-scale applications in water resource management. By defining hydrological objects as nodes and associating sensor-derived time-series data as node properties, the model offers a unified structure that seamlessly integrates temporal datasets. This approach is particularly effective in managing real-time data from sensors, which can be stored in time-series databases and linked to the graph data model for advanced analysis. The integration ensures comprehensive datasets that not only enhance the decision-making capabilities of the WMF and semantic web-based systems but also enable adaptive and sustainable management strategies.

Large-scale deployments, however, introduce challenges related to data integration, system compatibility, and scalability under high data loads. Distributed graph databases and middleware solutions can harmonize inputs from heterogeneous sources, ensuring seamless data integration. Future research should explore distributed architectures capable of supporting real-time updates, enabling the model to scale dynamically with evolving data demands. Additionally, incorporating advanced tools such as graph neural networks (GNNs) could uncover latent dependencies, providing deeper insights into hydraulic systems and advancing the boundaries of predictive analytics and decision support.

Expanding on its adaptability, future work should also focus on integrating the graph data model with real-time data sources, such as IoT devices, to enhance its responsiveness in dynamic environments. By merging real-time data warehousing with the graph structure, the model can support adaptive management strategies that respond to fluctuations in water availability, quality, and infrastructure conditions. Integrating climate and socioeconomic data into the model would further enable a holistic approach to water management, fostering comprehensive support for sustainable practices, resilient policy development, and optimized resource planning across diverse operational scenarios.

In conclusion, the proposed graph-based data model represents a transformative shift in water resource management, addressing the critical limitations of traditional frameworks. Its relationship-focused structure offers a robust foundation for dynamic, data-driven decision-making and complements established frameworks like the WMF and semantic web technologies by enhancing harmonization and reasoning capabilities. By incorporating real-time data integration and extending its scalability, the model paves the way for innovative, sustainable, and intelligent water resource management solutions. Future research will aim to refine these capabilities further, solidifying the model's role as a cornerstone of next-generation hydrological data systems.

6. Conclusions

This paper introduces a graph-based data model designed to address the complex, multi-dimensional relationships inherent in modern water resource management. By leveraging the flexibility of graph structures, this model overcomes the limitations of traditional relational databases, enabling a nuanced representation of intricate connections among hydraulic entities. Our framework comprehensively classifies and characterizes hydraulic objects, analyzes inter-object relationships, and constructs a multidimensional model spanning conceptual, logical, and physical layers. This model reveals both explicit and implicit relationships within hydraulic systems, significantly enhancing data integration, mining, and analytical capabilities. The result is a robust foundation that not only supports more precise and efficient water resource management but also strengthens decision-making in resource allocation and risk assessment, providing valuable insights for shaping water management policy and optimizing resource utilization.

Author Contributions: Conceptualization, J.X. and J.Q.; methodology, J.Q.; software, J.Q.; validation, W.X.; data acquisition, K.S. Shen; writing—original draft preparation, J.Q.; writing—review and editing, J.T. and Y.C.; supervision, J.X.; project administration, K.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Hubei Province, grant number 2023AFD202.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors Jiabin Qiao and Keyan Shen were employed by the company “Hubei Key Laboratory of Intelligent Yangtze and Hydroelectric Science, China Yangtze Power Co., Ltd.”. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

- Guo, Y.; Liu, L. Reservoir Control Operations and Water Resources Management. *Water* **2024**, *16*, 3000. [\[CrossRef\]](#)
- O’Connell, E. Towards adaptation of water resource systems to climatic and socio-economic change. *Water Resour. Manag.* **2017**, *31*, 2965–2984. [\[CrossRef\]](#)
- Xu, H.; Berres, A.; Liu, Y.; Allen-Dumas, M.R.; Sanyal, J. An overview of visualization and visual analytics applications in water resources management. *Environ. Model. Softw.* **2022**, *153*, 105396. [\[CrossRef\]](#)
- Kamyab, H.; Khademi, T.; Chelliapan, S.; SaberiKamarpashti, M.; Rezaia, S.; Yusuf, M.; Farajnezhad, M.; Abbas, M.; Jeon, B.H.; Ahn, Y. The latest innovative avenues for the utilization of artificial Intelligence and big data analytics in water resource management. *Results Eng.* **2023**, *16*, 101566. [\[CrossRef\]](#)
- Drogkoula, M.; Kokkinos, K.; Samaras, N. A comprehensive survey of machine learning methodologies with emphasis in water resources management. *Appl. Sci.* **2023**, *13*, 12147. [\[CrossRef\]](#)
- Deng, C.; Liu, L.; Li, H.; Peng, D.; Wu, Y.; Xia, H.; Zhang, Z.; Zhu, Q. A data-driven framework for spatiotemporal characteristics, complexity dynamics, and environmental risk evaluation of river water quality. *Sci. Total. Environ.* **2021**, *785*, 147134. [\[CrossRef\]](#) [\[PubMed\]](#)
- Behmel, S.; Damour, M.; Ludwig, R.; Rodriguez, M. Water quality monitoring strategies—A review and future perspectives. *Sci. Total. Environ.* **2016**, *571*, 1312–1329. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rehman, S.U.; Khan, A.U.; Fong, S. Graph mining: A survey of graph mining techniques. In Proceedings of the Seventh International Conference on Digital Information Management (ICDIM 2012), Macau, Macao, 22–24 August 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 88–92.
- Di Nardo, A.; Giudicianni, C.; Greco, R.; Herrera, M.; Santonastaso, G.F. Applications of graph spectral techniques to water distribution network management. *Water* **2018**, *10*, 45. [\[CrossRef\]](#)
- Yu, X.; Wu, Y.; Meng, F.; Zhou, X.; Liu, S.; Huang, Y.; Wu, X. A review of graph and complex network theory in water distribution networks: Mathematical foundation, application and prospects. *Water Res.* **2024**, *253*, 121238. [\[CrossRef\]](#) [\[PubMed\]](#)
- Howell, S.; Rezagui, Y.; Beach, T. Water utility decision support through the semantic web of things. *Environ. Model. Softw.* **2018**, *102*, 94–114. [\[CrossRef\]](#)
- Maidment, D.R. *Arc Hydro: GIS for Water Resources*; ESRI, Inc.: Redlands, CA, USA, 2002.
- Henshaw, D.L.; Sheldon, W.M.; Remillard, S.M.; Kotwica, K. ClimDB/hydroDB: A web harvester and data warehouse approach to building a cross-site climate and hydrology database. In Proceedings of the 7th International Conference on HydroScience and Engineering, Philadelphia, PA, USA, 10–13 September 2006.
- Daniel, E.B.; Camp, J.V.; LeBoeuf, E.J.; Penrod, J.R.; Dobbins, J.P.; Abkowitz, M.D. Watershed Modeling and its Applications: A State-of-the-Art Review. *Open Hydrol. J.* **2011**, *5*, 26–50. [\[CrossRef\]](#)
- Grolinger, K.; Hayes, M.; Higashino, W.A.; L’Heureux, A.; Allison, D.S.; Capretz, M.A. Challenges for MapReduce in Big Data. In Proceedings of the 2014 IEEE World Congress on Services, Anchorage, AK, USA, 27 June–2 July 2014; pp. 182–189.
- Almoradie, A.; Jonoski, A.; Popescu, I. Web Based Access to Water Related Data Using OGC WaterML 2.0. *Int. J. Adv. Comput. Sci. Appl.* **2013**, *3*, 83–89.
- Price, R. *Urban Hydroinformatics: Data, Models, and Decision Support for Integrated Urban Water Management*; IWA Publishing: London, UK, 2011.
- Katusiime, J.; Schütt, B. Integrated water resources management approaches to improve water resources governance. *Water* **2020**, *12*, 3424. [\[CrossRef\]](#)
- Angles, R.; Gutierrez, C. Survey of graph database models. *ACM Comput. Surv.* **2008**, *40*, 1–39. [\[CrossRef\]](#)
- Goleđzinowski, W.; Błocki, W. Social Network Analysis: From Graph Theory to Applications. *Soc. Commun.* **2023**, *1*, 151–164. [\[CrossRef\]](#)
- Pourhabibi, T.; Ong, K.L.; Kam, B.H.; Boo, Y.L. Fraud detection: A systematic literature review of graph-based anomaly detection approaches. *Decis. Support Syst.* **2020**, *133*, 113303. [\[CrossRef\]](#)
- Wang, S.; Hu, L.; Wang, Y.; He, X.; Sheng, Q.; Orgun, M.A.; Cao, L.; Ricci, F.; Yu, P.S. Graph Learning based Recommender Systems: A Review. In Proceedings of the International Joint Conference on Artificial Intelligence, Montreal, QC, Canada, 19–27 August 2021.
- Iklavov, Z.; Medvedev, D. Robust Reinforcement Learning on Graphs for Logistics optimization. *arXiv* **2022**, arXiv:2205.12888.

24. Jia, X.; Zwart, J.; Sadler, J.; Appling, A.; Oliver, S.; Markstrom, S.; Willard, J.; Xu, S.; Steinbach, M.; Read, J.; et al. Physics-guided recurrent graph model for predicting flow and temperature in river networks. In Proceedings of the 2021 SIAM International Conference on Data Mining (SDM), Virtual, 29 April–1 May 2021; SIAM: Philadelphia, PA, USA, 2021; pp. 612–620.
25. Ponti, A.; Candelieri, A.; Giordani, I.; Archetti, F. A novel graph-based vulnerability metric in urban network infrastructures: The case of water distribution networks. *Water* **2021**, *13*, 1502. [[CrossRef](#)]
26. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Yu, P.S. A Comprehensive Survey on Graph Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 4–24. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.