An Effective Digital Twin Modeling Method for Infrastructure: Application to Smart Pumping Stations

Fan Feng, Zhansheng Liu *, Guoliang Shi and Yanchi Mo

Faculty of Architecture, Civil and Transportation Engineering, Beijing University of Technology, Beijing 100124, China; fengemma0515@gmail.com (F.F.); guoliangshi@emails.bjut.edu.cn (G.S.); yanchimo@emails.bjut.edu.cn (Y.M.)
* Correspondence: liuzhansheng@bjut.edu.cn

Abstract: Digital twin technology has evolved from a theoretical concept to practical application, facilitating seamless data exchange between virtual and physical domains. Although there has been progress, the infrastructure industry, which is recognized for its intricate nature and the need for timely action, is still in the first phases of digital twin advancement. A significant obstacle in this field is the absence of established definitions and modeling standards, which impede the precise depiction of infrastructure systems. To address these challenges, this paper proposes a high-precision digital twin modeling method tailored for pumping stations. The method focuses on two key scenarios: first, we construct an overall digital twin model that contains both physical entities and operational processes of pumping stations; second, we design a modeling process applicable to pumping stations by analyzing the deficiencies of the existing standard system. Additionally, we selected the East–West Water Transfer Project in China as a case study to demonstrate the high-precision digital twin model of a pumping station. This model will include essential components, such as the modeling of pumping stations, the operational processes of pumping stations, and the modeling of system operation analysis. Serving as the database for the digital twin, it can complete the automatic inspection of the pumping station, optimization of scheduling, prediction and regulation of energy and carbon emissions, and visualization of results for display and other applications. The model realized the benefits of 100% automatic inspection rate, reduction of eight corresponding operating personnel, and comprehensive cost saving of RMB 2.25 million. The objective of this research is to narrow the divide between theoretical concepts and real-world implementations by pushing the boundaries of digital twin technology in the specific context of intricate infrastructure projects. This project aims to improve the practicality of digital twin technology in real-world situations, namely in the infrastructure industry.

Keywords: effective modeling; pumping station; digital twin; method research

1. Introduction

In 2003, Michael Grieves introduced the digital twin concept. NASA later adopted this concept for aircraft design, marking a shift towards the incorporation of virtual reality and real-time engagement in global manufacturing [1]. Subsequently, Tao Fei et al. [2] improved upon Grieves’ original three-dimensional framework by developing a five-dimensional digital twin model. This work aimed to align with the changing trends and increasing demands in the sector, indicating a notable progression in digital twin technology.

Tao Fei et al. [3] conducted a comprehensive analysis on the concept, enabling technologies and varied applications of digital twins, and proposed ten pivotal questions to enhance the overall understanding of this technology. The objective of this investigation was to enhance the comprehension of digital twins and their potential in various dimensions and sectors. The increasing implementation of digital twin technology in several sectors has...
led to a greater need for accuracy in twin modeling. Hence, crucial concerns such as the
accuracy of virtual models compared to their real-life counterparts as well as their capacity
for monitoring, modeling, forecasting, and optimizing physical entity operations have been
the primary focus of extensive scholarly research.

In recent years, scholars have used digital twin technology to manage construction and
operation in the infrastructure domain. Notable examples include Ramos Helena M et al. [4],
who developed a smart water grid (SWG) for water infrastructure using digital twins, which
enhanced system efficiency and management capabilities. Similarly, Haohan T et al. [5] in-
troduced a digital twin-based fault diagnosis framework for data-driven underwater control
systems, employing a novel Physical-Informed Time Domain Convolutional Network (PITCN)
that surpasses traditional CNN-based methods. Additionally, Shengwen Zhou et al. [6] uti-
lized the five-dimensional model to effectively represent a water purification plant at
multiple levels, demonstrating its practical application.

Infrastructure projects stand out due to their complexity, variability, extensive scale,
and prolonged durations, presenting unique challenges for digital twins. However, data on
Infrastructure Digital Twins (IDTs) in this sector are limited and there is still no consensus
on its definition. As a result, integrating data, models, and services in digital twin model
construction faces significant challenges. With the intention of merging the specific charac-
teristics of the infrastructure domain with high-precision digital twin modeling, this study
seeks to provide a substantial reference for future research. The primary research questions
addressed are as follows:

- **RQ1:** What are the fundamental processes for constructing an intelligent digital twin
  model for the infrastructure domain?
- **RQ2:** What are the current digital twin standard systems in place?
- **RQ3:** What are the principal techniques for creating IDTs?
- **RQ4:** How to develop high-precision digital twin models for the infrastructure domain?

This paper presents the key processes and methods for developing high-precision
digital twin models in the infrastructure sector, drawing from a thorough literature review
and international standards. A real project case is then chosen to evaluate the practicality
and precision of these methods.

Section 2 will describe the research methodology, including the steps taken to create
twin models that take into account infrastructure characteristics, and addresses RQ1. In
Section 3, we will discuss the findings related to standards and technologies. These findings
will help develop a framework for digital twin modeling in the infrastructure domain,
addressing both RQ2 and RQ3. Section 4 will focus on the evaluation of the proposed
modeling approach and technical system through a real project case. Lastly, Section 5 will
provide a summary of this study’s findings.

2. Research Methodology

2.1. Basic Process for Building Models

To ensure seamless interconnectivity of digital twin systems and to minimize conflicts
across different domains, it is imperative to establish clear and consistent principles and
processes for constructing digital twin models. The digital twin model, serving as a
virtual representation of physical entities or systems, is central to overseeing, predicting,
optimizing, and managing real-world situations. The precise and thorough creation of a
digital twin model is crucial for enabling seamless integration between virtual and real
environments and promoting a mutually beneficial connection. When creating a digital twin
model, it is crucial to take into account not just the physical entity, but also its operational
procedures, traits, methodologies, behaviors, and other pertinent aspects.

The methodology employed in this research aligns with the five-dimensional model
theory of the digital twin by Tao Fei et al., as outlined in their work [7]. During model
construction, our approach considers the “geometry-physics-behavior-rule” dimensions
to ensure a comprehensive representation. Through integrating various fields including
mechanical, electrical, fluid dynamics, and topology, our methodology enables the mod-
eling of basic units, complex objects, and spatial systems. Consequently, it provides a detailed representation of complex physical entities in the real world. Figure 1 shows the technical roadmap implemented in this research, which is organized into four main components: physical entity modeling, virtual entity modeling, decision model modeling, and visualization application. The process begins with sensing the physical entity using sensors, followed by creating a virtual model through parametric modeling and multidimensional data. Establishing the connection between the physical entity and the virtual model is achieved through the method of virtual-reality map. Subsequently, an appropriate algorithm is selected to create a decision model, which is based on the integration of data fusion and model coupling. Finally, the outcomes of data analysis are presented in a visual format for enhanced interpretation.

![Figure 1](image1.png)

**Figure 1.** The technological roadmap of this study.

Effective model management is necessitated by the process of construction, verification, and correction, as well as subsequent utilization to guarantee the model’s accuracy and validity within the twin system. It is imperative to validate the model’s realism post-construction, and if tests reveal inaccuracies, the model undergoes refinement through a “verification-correction” cycle. This cycle is repeated as necessary until the model meets the required accuracy standards.

We have developed a basic model construction process depicted in Figure 2, based on the characteristics of the infrastructure lifecycle construction stage. Throughout the lifecycle from design to construction to operation, real-time mapping relationships exist between physical and virtual entities. The process involves model building, assembling, fusion, validation, correction, and management to ensure the accuracy of the model.

![Figure 2](image2.png)

**Figure 2.** Basic process for building digital twin models in infrastructure domains.
2.2. Basic Principles for Building Models

To standardize the construction process of the digital twin model, a modeling criterion for the digital twin is proposed, based on previous research on the five-dimensional model. Additionally, specific principles governing the building process of the digital twin model are outlined to align with modeling requirements and intended digital application purposes, as illustrated in Figure 3.

![Figure 3. Digital twin modeling principles.](image)

The fundamental purposes of the principles of digital twin modeling, as depicted in the figure, include “usable, versatile, efficiency, user-friendly, connectivity, integrated, flexible and intelligent”. In addition, the model must also fulfill the requirements of “precision, standardization, lightweight and visualization” to ensure model refinement. Precision entails the accurate and consistent static portrayal or dynamic result output of the entity at the physical or system level, ensuring that the result is credible and usable, meeting the validity requirements of the model. Standardization involves the establishment of a unified standard for the construction of various elements and objects in different industries and domains, ensuring the consistency of model integration, data exchange, information identification, and maintenance, thereby enhancing the model’s versatility. Lightweight refers to streamlining model information and structure without compromising main information, accuracy, and functionality, ultimately improving operational efficiency. Finally, visualization involves presenting the entire lifecycle activities of the model in a more intuitive manner, facilitating deeper interaction with the user and enhancing the model’s readability and usability.

The digital twin model should be designed to meet the demands of various application scenarios, including being “reconfigurable, integrable, interactive and evolutionary”. Reconfigurability entails the model’s ability to adapt to complex and changing application environments by flexibly changing its structure, composition, parameter configuration, and relationships. Similarly, integrability necessitates the ability to line multiple digital twin models and integrate data from various scales and dimensions to aid in system operation and decision-making. Interactive capabilities involve enabling connectivity, breaking information silos between systems, and ensuring interoperability between entities, models, and data. Lastly, meeting evolutionary demands requires the model to be intelligent, constantly evolving following the operating principles of the physical entity or system, and continuously optimizing and correcting itself based on historical and dynamic process data.

2.3. Theoretical for Building Models

The digital twin theory, which is applied from the manufacturing field to the architectural field, aims to address the challenge of “information-physical” non-interaction by focusing on the “physical” building transferring its real status to the “information”.

However, the establishment of digital twin models is confronted with issues such as system complexity, uncertain external environment, and variable coupling. Therefore, to achieve a high-precision and accurate reflection of the real characteristics, it is essential to address the coupling of different physical and time scales. As noted by references, multi-scale modeling is vital for linking information from different time and physical scales to scientifically simulate buildings in the infrastructure field and create a high-precision twin model. Building on this, this paper introduces the all-element digital twin model $DTM_p$ for pumping stations, as depicted in Figure 4. The model is represented by Equation (1).

$$DTM_p = \{GM_p, PM_p, DM_p\}$$ (1)

![Figure 4. All-element digital twin model for pumping station.](image)

The geometric model ($GM_p$) in the digital twin model of a pumping station is a true-to-life representation of the physical entity in a digital space, encompassing its geometric dimensions, shapes, positions, materials, and environmental characteristics. The geometric model is expressed as Equation (2).

$$GM_p = \{G_S, G_G, G_P, G_M, G_E\}$$ (2)

where $G_S$ is size data; $G_G$ is shape data; $G_P$ is position data; $G_M$ is material data; $G_E$ is environment data.

The process dynamic model ($PM_p$) in the pumping station digital twin model serves as a real-time representation of the operating state of physical entities in a digital space. This facilitates information connection and transmission among equipment, systems, and their respective components, using methods such as real–virtual interaction mapping. Key components of the model include descriptions of the attribute information of physical objects (e.g., data, interfaces, states, relationships), as well as the state and cooperative work...
of the internal components and between the objects. The process model is expressed as Equation (3).

\[ PM_P = \{P_{ID}, P_P, P_A, P_R, P_S\} \tag{3} \]

where \( P_{ID} \) is the identification of the physical object; \( P_P \) is the attribute describing the physical object, including static and dynamic attributes; \( P_A \) is the operation of the physical object; \( P_R \) is the rule of the physical object; \( P_S \) is the subcomponent of the physical object (if any).

The decision model (\( DM_P \)) in the pumping station digital twin model integrates historical and real-time data of physical entities when they operate in the digital space, thus enabling the prediction and feedback on the operation status and results. These results are ultimately presented visually to the application. With decision models, there is a combination of data-based smart models, mechanistic models, and hybrid models integrating both smart and mechanistic elements. The decision model is expressed as Equation (4).

\[ DM_P = \{M_S, M_M, M_SM\} \tag{4} \]

where \( M_S \) is a data-based smart model; \( M_M \) is a mechanistic model; \( M_SM \) is a mixture of an intelligent model and a mechanistic model.

3. Literature Review

3.1. Digital Twin Modeling Standard System

Our investigation revealed the existence of both domestic and international standards that provide specific guidelines and technology for each component of the digital twin five-dimensional model. Adhering to these standards is essential at the stage of constructing the refined digital twin model. This research thoroughly examines the recent standards relevant to digital twin models in the building and infrastructure domains, organizing them into categories such as physical entities, virtual entities, data, and connections. This classification has led to the development of a precise standard system for modeling digital twins, as depicted in Figure 5.

![Digital twin model for infrastructure domains](image)

**Figure 5.** Standards related to digital twins in the building/infrastructure domains [8–26].

Our study identified various international and domestic standards relevant to the construction of the digital twin five-dimensional model. These standards provide guidance for refining the model at each stage of the construction process.

The realm of digital twins involves the use of physical entities to represent real-world objects for the purpose of modeling, simulation, and analysis. Different domains have their own tailored definitions; for example, in manufacturing, the focus is on the production equipment (ISO 23247) [8]. In architecture, the emphasis is on the geometric, spatial, functional, and performance attributes of buildings (ISO 16739) [9], while urban planning incorporates the relationships between city infrastructure and GIS data (CityGML) [27–29].
Virtual entities, on the other hand, are multi-dimensional, multi-scale digitized versions of their physical counterparts. Standards such as ASTM E3012-16 [10] and ISO 16739 [9] provide modeling frameworks that cover the design, operations, and maintenance phases, while ISO 19650 specifically focuses on the full lifecycle management of building and infrastructure virtual models [11].

The functionality and efficacy of digital twins depend fundamentally on the data. Standards play a crucial role in enabling interoperability and data exchange. Notably, standards such as IFC for Building Information Modeling software [11], STEP for product data exchange [13,30], and CityGML for virtual city model storage and exchange have been developed for this purpose [16]. Moreover, ISO/IEC 30182 [14] provides a data sharing framework that supports collaborative analysis of city services. In addition, data security and access control [15] and ISO/IEC 27001 [17] for information security management are essential for ensuring the integrity and privacy of data involved. Furthermore, connectivity and communication protocols, such as OPC Unified Architecture (OPC UA) [31], MTConnect [32,33], Modbus, and BACnet [34,35], are crucial for facilitating precise data flow and interoperability within the industry. Semantically, standards like RDF and OWL are important for semantic mapping [36], while IFC and STEP are foundational for data conversion. This demonstrates the pivotal role that connectivity and communication protocols play in facilitating information exchange among entities in the digital twin ecosystem.

Digital twins can significantly impact real-world applications and services, particularly in the developing domains of building and infrastructure. Although standards in these areas are still evolving, there are notable examples such as ISO 19650 for project and asset management, BS 1192-4:2014 [19] for information exchange protocols, and ISO standards for smart city and transportation timing guidelines [20,21]. These service-focused standards collectively establish a framework for specification and implementation, covering various aspects including description, development, deployment, operation, and testing. Key standards in this context include ISO 13372:2012 [22], ISO 17359:2018 [23], ISO/IEC 12207 [24], the ITIL Framework, IEEE 1232.3-2014 [25], and ISO 20242-3-2011 [26].

3.2. Digital Twin Accurate Modeling Technology System

This paper further elaborates on the technical systems required for the accurate construction of digital twin models in infrastructure, encompassing the entire lifespan and varying technological types, as depicted in Figure 6. An Infrastructure Digital Twin (IDT) is conceptualized as a digital replica that precisely mirrors a physical asset or entity, facilitating bidirectional data exchange between the virtual and physical realms. Achieving such precision necessitates meticulous geometric modeling and a framework for real-time, realistic feedback. This system can operate autonomously or in conjunction with other systems, offering a more comprehensive representation of infrastructure operations and thereby enhancing its performance. Essential to this endeavor is the gathering and integration of data, understanding component interrelationships, and ensuring accessible data for the twin system. Building on the work of Tao Fei et al. [7] and Naderi H et al. [37] and considering the unique attributes of the infrastructure sector, the focus is placed on the detailed technical aspects necessary for the development of accurate digital twin models.

The critical data acquisition technologies enabling communication between physical and virtual entities, such as IoT sensors [38], point cloud scanning methods [39], RFID, QR codes, and SCADA systems [40], are identified in the existing literature for operational and maintenance problem-solving. Information modeling primarily relies on BIM, GIS, simulation, and finite element analysis [41]. BIM provides essential system detail; on the other hand, GIS demonstrates its capability in large-scale geospatial infrastructure modeling. Additionally, simulation and emerging technologies such as VR, AR, and robotics show potential in expansive scenario analysis, as noted by scholars.
Figure 6. A technical system for building digital twin models in infrastructure.

Academics observe a significant rise in the collection of infrastructural data in recent times. Nevertheless, despite the rapid expansion, the availability of effective processing technologies has been somewhat restricted, but they are increasingly being adopted in diverse fields such as machine learning, big data analytics, computer vision, and algorithm optimization. The emergence of cloud computing and edge computing has also contributed to advancement in this area. In terms of interoperability, several solutions have been proposed in the literature to integrate IoT sensor data into Building Information Modeling (BIM). Notable methods include the use of IfcSensor [42], the Revit API [43], or SensorML [44], all of which have demonstrated analytical success. Additionally, geospatial application analysis has relied on IfcOpenShell [45] and sought to facilitate BIM/GIS data exchange via shapefiles [46]. Moreover, the integration of data and semantic rules has played a crucial role in facilitating schema transformation across various frameworks such as Industry Foundation Classes (IFC), City Geography Markup Language (CityGML), and ontologies [47,48]. This integration has sparked significant scholarly interest due to its potential to elicit substantial advancement in the field.

3.3. Current State and Challenges

In recent years, digital twin technology has been widely recognized and applied in many fields, with the infrastructure sector being no exception. However, research to date has predominantly focused on investigating the conceptual framework, general modeling techniques, and use cases of digital twins. In contrast, there is a notable lack of in-depth studies on digital twin modeling methodology specifically tailored for the infrastructure industry.

In terms of conceptual framework, Tao et al. [2] proposed the widely accepted five-dimensional digital twin modeling theory. Lu et al. [49] focused on digital twin-driven smart manufacturing within the context of Industry 4.0. They presented a digital twin architecture model, examined current status and progress, outlined existing digital twin applications and typical scenarios, and discussed the key issues for future research.

In terms of general modeling techniques, Kritzinger et al. [50] reviewed the most important techniques for digital twin modeling, such as virtual reality, augmented reality, big data, etc. Zhou et al. [51] proposed a generic framework for knowledge-driven digital twin manufacturing units for smart manufacturing, which can support autonomous manufacturing. However, it is often difficult to fully address the unique complexities and limitations of these generic approaches when applied to specific domains, such as the infrastructure industry.

In view of this, scholars have begun to focus on infrastructure digital twin technology. Jiang et al. [52] proposed a digital twin model for information management of complex
infrastructure systems to provide a unified reference specification and description methodology for information management. Mohammadi et al. [53] integrated a TSL-driven Bridge Information Model (BrIM) with a Decision Support System (DSS) approach to realize bridge infrastructure with digital twin and asset management. However, most of the previous studies concentrate on applying a twin model in a particular scenario, but they lack detailed research on the methodology of building high-precision models and on the overall process of modeling a digital twin.

Compared to the existing studies, the innovations of this paper mainly include the following:

1. This study aims to address the digital twin problem in modeling complex infrastructure systems and proposes a high-precision modeling method based on general modeling techniques.
2. According to the five-dimensional theory and infrastructure characteristics, a novel high-precision modeling process is introduced, along with a systematic technical framework.
3. The modeling methodology is validated empirically using real infrastructure projects as cases and assessing their quantitative benefits.
4. The main implementation challenges are analyzed in detail and effective solutions are identified to serve as a reference for later promotion.

In summary, this study emphasizes the real needs of the infrastructure industry and proposes novel ideas for high-precision approaches that are of great theoretical significance and application value.

4. Case Study

The urban water system is a crucial element in urban infrastructure, providing essential support for urban life. However, it is a highly intricate system, with the majority of its components located underground. Monitoring its operation process is challenging, and repairing any identified issues poses additional difficulties. Pumping station as the core part of the urban water supply system is an important infrastructure related to the survival and development of the city [54]. Pumping stations also serve as the main means to realize inter-basin water transfers from water-rich areas to water-scarce areas in other basins to replenish water sources. Compared with other infrastructure projects, pumping stations are characterized by a high degree of automation, high equipment complexity, and high failure costs.

As an important part of the water conservancy project, the operation of the pumping station is closely related to water resources’ scheduling and distribution, and the monitoring and regulating function of the pumping station on water quantity and quality is extremely critical. In the process of its operation, the hydromechanical parameters change in a wide range and the equipment has a great impact, low fault tolerance, and the need for fine monitoring. Due to the need for unattended and remote monitoring, the pumping stations rely more on automated operation and intelligent control, and the requirements for digitalization and informatization are more stringent. In addition, large pumping stations not only have numerous pumping equipment, but also have complicated pipe network systems, gate and dam systems, electrical control systems, and complex composition. Pumping station downtime may threaten the safety of the entire water conservancy project water supply, causing serious consequences, so the operation and maintenance management requirements for pumping stations are also extremely high. Introducing the concept of digital twin and constructing an accurate and realistic digital twin model, which can reflect the asset performance in real time, help facility managers to evaluate the data fed back from the twin system, judge the operation situation, and analyze the high-energy parts, so as to fully control the operation of the facility; in addition, it is possible to simulate the extreme situation through the digital twin system for failure analysis, identify the potential risks, and provide early warning.
4.1. Case Presentation and Objectives

The objective of the East–West Water Transfer Project in China is to mitigate the disparity in land resources between the east and west regions. The east region possesses abundant water resources but limited land, while the west region has ample land but scarce water resources. For this study, we selected the East–West Water Transfer Project in Beijing as our subject, with a particular emphasis on the Xingshikou Pumping Station. The objective is to demonstrate the construction of a high-precision model using digital twin technology to assist in the intelligent monitoring of the pumping station’s operation throughout the entire process.

The pumping station is composed of various rooms such as the central control room, main plant, low-voltage room, inverter room, high-voltage room, automation room, and other equipment rooms. These rooms house different equipment including distribution boxes, flow meters, high-pressure and low-voltage distribution cabinets, inverter cabinets, pump bodies, pump motors, pipelines, and other key equipment, as presented in Figure 7.

![Figure 7. Layout of pumping station.](image)

4.2. Data Collection

In this project, in order to achieve comprehensive and uninterrupted data collection that aligns with the basic requirements for healthy operation and daily inspection of the pumping station, it was necessary to address the inadequacies of the number and location of the existing sensors. The initial sensors in place were unable to accurately capture the actual working conditions and failed to provide sufficient data to fulfill the management needs of the intelligent pumping station. Consequently, we undertook the selection, development, and deployment of new sensors in conjunction with the existing ones. The configuration list of sensors used in this project is shown in Table 1.

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>Position</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vibration sensors</td>
<td>3 motor vibration, 3 pump vibration</td>
<td>72</td>
</tr>
<tr>
<td>2</td>
<td>Oscillation sensor</td>
<td>Mounted on the main shaft of the unit</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>RPM (revolutions per minute) sensor</td>
<td>Mounted on the main shaft of the unit</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>Noise sensor</td>
<td>Installation next to the machine (1~2 m)</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>Signal amplifier</td>
<td>Mounted in the system cabinet</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>Position</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Industrial switches</td>
<td>Mounted in the system cabinet</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>High-precision data acquisition system (16 channels)</td>
<td>Mounted in the system cabinet</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>System server</td>
<td>Mounted in the system cabinet</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>Vibration analysis software module</td>
<td>Mounted in the system cabinet</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>Temperature analysis software module</td>
<td>Mounted in the system cabinet</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>Radar water level meter</td>
<td>Forebay or catch basin</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>Temperature and humidity sensors</td>
<td>Inspection room</td>
<td>21</td>
</tr>
<tr>
<td>13</td>
<td>Unit temperature sensors</td>
<td>Xingshikou pumping station unit</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>Unit temperature transmitter</td>
<td>Xingshikou pumping station unit</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>Centralized power supply</td>
<td>Computer room</td>
<td>7</td>
</tr>
<tr>
<td>16</td>
<td>Transformer temperature controller</td>
<td>Inside the transformer in the</td>
<td>6</td>
</tr>
</tbody>
</table>

The configured sensors are used according to the data acquisition requirements, and the application contents and locations are listed in Table 2.

Table 2. Sensor deployment.

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Content</th>
<th>Quantity</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vibration sensors</td>
<td>Motor vibration</td>
<td>The generator housing is horizontally, vertically, and radially welded to the metal base.</td>
<td>3</td>
<td>Depending on the conditions on site, insulated epoxy resin plinths can also be installed. Install the device in a suitable location depending on the maintenance requirements at the site.</td>
</tr>
<tr>
<td></td>
<td>Water pump vibration</td>
<td>The water pump surface is horizontally, vertically, and radially welded to the metal base.</td>
<td>3</td>
<td>Install the device in a suitable location depending on the maintenance requirements at the site.</td>
</tr>
<tr>
<td>Oscillation sensor</td>
<td>X-direction oscillation</td>
<td>Non-contact measurement of the major axis oscillation and a U-bracket welded to the bottom of the major axis.</td>
<td>1</td>
<td>Install the device in a suitable location depending on the maintenance requirements at the site.</td>
</tr>
<tr>
<td></td>
<td>Y-direction oscillation</td>
<td>Non-contact measurement of the major axis oscillation and a U-bracket welded to the bottom of the major axis.</td>
<td>1</td>
<td>Install the device in a suitable location depending on the maintenance requirements at the site.</td>
</tr>
<tr>
<td>Temperature sensor</td>
<td>Motor temperature</td>
<td>Generator housing horizontal, vertical, and radial welded or affixed pressure plate temperature sensors. Water pump surface horizontal, vertical, and radial welding or adhesion of pressure plate temperature sensor.</td>
<td>3</td>
<td>Install the device in a suitable location depending on the maintenance requirements at the site.</td>
</tr>
<tr>
<td>Pump temperature</td>
<td></td>
<td>Weld a U-bracket to the bottom of the large shaft (can be shared with the oscillation bracket).</td>
<td>3</td>
<td>Install the device in a suitable location depending on the maintenance requirements at the site.</td>
</tr>
<tr>
<td>RPM sensor</td>
<td>Rotation speed of the unit</td>
<td>1.5 cm diameter reflective sheet needs to be attached to the main shaft.</td>
<td>1</td>
<td>Install the device in a suitable location depending on the maintenance requirements at the site.</td>
</tr>
<tr>
<td>Noise sensor</td>
<td>Noise of the unit</td>
<td>Concealed installation</td>
<td>1</td>
<td>Install the device in a suitable location depending on the maintenance requirements at the site.</td>
</tr>
</tbody>
</table>
A wide variety of precise sensor systems have been constructed by replacing and expanding the current sensor equipment. This advancement addresses the inadequacies of existing sensors that fail to meet actual requirements and lack sufficient capabilities in monitoring equipment operating status. Through signal transmission, the sensors can retrieve information about the status of the devices in real time and perform automatic inspection. The sensor architecture is shown in Figure 8.

![Sensor Architecture](image)

**Figure 8.** The sensor architecture.

The data collected in this project include real-time monitoring data from sensors, historical data from pumping station operations, and data provided by government departments, as shown in Table 3.

**Table 3.** Data collection content.

<table>
<thead>
<tr>
<th>No.</th>
<th>Category</th>
<th>Type</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Real-time data collection by sensors</td>
<td>Structured</td>
<td>Real-time</td>
</tr>
<tr>
<td>2</td>
<td>East–West Water Transfer statement data</td>
<td>Structured</td>
<td>Once gathering</td>
</tr>
<tr>
<td>3</td>
<td>Water conditions in large- and medium-sized reservoirs</td>
<td>Structured</td>
<td>Real-time</td>
</tr>
<tr>
<td>4</td>
<td>Urban rainfall data</td>
<td>Structured</td>
<td>Real-time</td>
</tr>
<tr>
<td>5</td>
<td>Radar cloud map data</td>
<td>Unstructured</td>
<td>Irregular</td>
</tr>
<tr>
<td>6</td>
<td>Flood warning and forecasting</td>
<td>Unstructured</td>
<td>Irregular</td>
</tr>
<tr>
<td>7</td>
<td>Maintenance and repair data</td>
<td>Unstructured</td>
<td>Irregular</td>
</tr>
<tr>
<td>8</td>
<td>Basic information data of the project</td>
<td>Structured</td>
<td>Irregular</td>
</tr>
<tr>
<td>9</td>
<td>Power monitoring data</td>
<td>Structured</td>
<td>Real-time</td>
</tr>
<tr>
<td>10</td>
<td>Power fail-safe data</td>
<td>Structured</td>
<td>Real-time</td>
</tr>
</tbody>
</table>

The data collected by the sensor are listed in Table 4. The real-time data are transmitted to the internal database of the East–West Water Transfer Project through the data collection platform, and other types of data are collected according to the scope and update frequency in the database, which facilitates access to the system.

Data acquisition and processing play a pivotal role in determining model accuracy and influencing decision-making. This process encompasses data integration and analysis, linking equipment condition databases with computer systems via BIM models to ensure comprehensive recording of maintenance data. Utilizing BIM management concepts, data and asset models are formulated, with IFC architecture and COBie formats, facilitating...
data transfer. The architecture for collecting and processing data through IoT sensors is shown in Figure 9.

### Table 4. Data collection from sensors.

<table>
<thead>
<tr>
<th>No.</th>
<th>Equipment</th>
<th>Data Category 1</th>
<th>Data Category 2</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Main water pump</td>
<td>Operational monitoring</td>
<td></td>
<td>Vibration</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>Vibration Y</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>Oscillation</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>Oscillation X</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>Thrust axis temperature</td>
<td>Thrust axis temperature</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>Outlet pressure</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>of main pump</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>Reactive power Q</td>
</tr>
<tr>
<td>9</td>
<td>Auxiliary drive</td>
<td></td>
<td></td>
<td>Active power P</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td>Power factor cos</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td>Stator winding temperature 1</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td>Stator winding temperature 2</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td>Stator winding temperature 3</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td>Stator winding temperature 4</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td>Stator winding temperature 5</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td>Stator winding temperature 6</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td>High-voltage soft start</td>
<td>Motor current Ia</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td>current</td>
<td>Motor current Ib</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td>Motor current Ic</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td>High-voltage soft start</td>
<td>Motor voltage Uac</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td></td>
<td>line voltage</td>
<td>Motor voltage Uab</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td>Inverter ambient temperature</td>
</tr>
<tr>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td>Inverter winding temperature</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td>U</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td>Inverter winding temperature</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td>V</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td>Inverter winding temperature</td>
</tr>
<tr>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td>W</td>
</tr>
<tr>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td>Inverter frequency setting</td>
</tr>
<tr>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td>Rate</td>
</tr>
<tr>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td>Frequency</td>
</tr>
<tr>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td>Current</td>
</tr>
<tr>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td>Input voltage</td>
</tr>
<tr>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td>Power (output)</td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td>Output flow rate</td>
</tr>
<tr>
<td>41</td>
<td></td>
<td></td>
<td></td>
<td>Forebay water level</td>
</tr>
<tr>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td>Catch basin level</td>
</tr>
</tbody>
</table>

The data analysis layer uses advanced machine learning techniques. For example, it uses deep learning to process and interpret large-scale data. This analysis achieves anomaly detection, fault diagnosis, performance prediction, and energy consumption assessment. The insights are visually presented to project managers. This helps them make informed decisions for continuous operational and maintenance improvements.

### 4.3. The Method Framework

We have developed a digital twin architecture for an intelligent pumping station to achieve precise operational control. This architecture, shown in Figure 10, consists of five components: the physical pumping station, its digital twin, connection interfaces, twin data, and an intelligent service layer. The physical station comprises the main plant, the electromechanical equipment, and the regulating reservoirs. The digital twin accurately reflects the actual production activities in the real world. The connection interface enables bidirectional data exchange and integration between the physical and virtual worlds. The
twin data act as a link between these worlds and aggregate operational and simulation data, which are essential for predictive analysis. The intelligent service layer, which uses time series decomposition (STL) and long short-term memory networks (LSTM), monitors and analyzes key parameters such as carbon emissions, energy consumption, and water demand. This leads to improved operational efficiency of the pumping station and reduces energy consumption.

![Figure 9. Framework of data acquisition and processing.](image)

![Figure 10. Pumping station digital twin system architecture.](image)

### 4.4. Modeling Development

In Section 4.2, the physical model data collection framework has been established, with defined procedures for data collection. Following this, in Section 4.3, the digital twin five-dimensional model architecture of the pumping station has been constructed in accordance with the technical process. This section focuses on detailing the model development process, encompassing key areas such as geometric model construction, data
and model assembly and fusion, verification of consistency between the virtual and real models, as well as the establishment of a digital twin operating system for the intelligent pumping station based on the high-precision model. The development process of the virtual model is shown in Figure 11.

Figure 11. Virtual modeling development process.

(1) The geometric model of pump station ($GM_P$)

In scenarios where CAD drawings are not available, the initial step in the modeling process is to break down complex spatial data into fundamental components representing "man, machine, material, method, and environment". These detailed data are used to create a geometric model of the building using specialized modeling software. The model is designed to incorporate real-world project constraints and is developed with varying levels of accuracy, as illustrated in the technology roadmap presented in Figure 12.

At the beginning of the workflow, building contours, roof lines, and wall surfaces are first extracted from the aerial image or DOM data using digital photogrammetry. This step lays the foundation for automatic building modeling. Subsequently, the texture data captured by UAVs and ground photography are analyzed and processed to remove obstacles such as angular distortions and occlusions and to ensure precise alignment with the building models.

The method consists of extracting spatial object contours from topographic maps and converting them into compatible formats for modeling software to enable the creation of detailed models. The creation of 3D models is performed manually using organized photos and additional data such as CAD and aerial imagery. These data are used to classify spatial objects and construct models with varying degrees of precision.

The geometric model of the pumping station is created at multiple scales, and the visualization model is divided into three levels: macro (S1), meso (S2), and micro (S3). At the macro level, the modeling includes the overall scene of the Xingshikou plant. The meso level includes the representation of the mechanical and electrical equipment, such as the main plant, starters, gates, high- and low-voltage rooms, etc. This level aims to visualize and simulate the structure and operating conditions of the equipment of the water conservancy project. The micro level includes the low-voltage distribution room, which
consists of two sets of low-voltage power distributors, distribution boxes, and flow meters. The high-voltage distribution room includes 14 high-voltage distribution cabinets and a comprehensive DC protection cabinet. The frequency converter room contains four sets of frequency converter cabinets, while the automation room houses automation control cabinets. Finally, the central control room has, among other things, a command screen, as shown in Figure 13.

![Diagram of data extraction technology roadmap](image)

**Figure 12.** Buildings’ data extraction technology roadmap.

![Diagram of pumping station model](image)

**Figure 13.** Pumping station model.
(2) The process dynamic model of pump station ($PM_P$)

In the area of digital twin models for infrastructures, the focus goes beyond the analysis of correlations between physical entities and system operations to include simulations and optimizations of system behavior, performance, and maintenance approaches. Merging and fusing multiscale models that encompass all elements means that unit-level models are integrated into cohesive system-level representations, considering spatial and constraining relationships while considering operational procedures, maintenance strategies, and environmental impacts. The goal is to comprehensively simulate the performance of the entire infrastructure system, considering not only the physical and mechanical aspects but also the operational logic, environmental adaptability, and long-term maintenance aspects. An effective fusion of models should enable predictive analysis under different scenarios and support the formulation of scientific and sustainable management strategies.

By integrating data from different sources of the pumping station using the previously discussed data collection and fusion methods, we combined different types of data, including basic data, business data, and heterogeneous data. To improve the speed of scene visualization and enable real-time interaction, we implemented the display of details in different resolutions and perspectives. The Water Authority’s electromechanical devices, such as pumps and sluices, are stored in either .max or .ifc format to support streaming and rendering of complicated model data. This format provides a dynamic representation of the spatial location and operational status of the devices, as can be seen in Figure 14, which shows a highly detailed model of a water pump.

![Figure 14. (a) Real pumps; (b) high-precision pumping station.](image)

(3) Model Verification

Consistency checking in digital twin modeling is critical to ensure consistency between the physical entity and its digital representation in four dimensions: geometric, physical, behavioral, and rule-based.

Geometric dimension: Precise measurement techniques such as laser scanning and photogrammetry are used to obtain geometric structural data. These data are then compared with the modeled data using tolerance analysis to confirm geometric consistency.

Physical dimension: During validation, the measured data are compared with the model predictions to confirm the accuracy of the physical model.

Behavioral dimension: Simulation of system operation using time or state diagrams is used to verify the model’s accurate representation of system behavior.

Rule dimension: Comparison of the system response predicted by the model with the actual operational responses to validate the rule-based modeling.

(4) The decision model of pump station ($DM_P$) and results visualization

In this study, we have developed simulation models that focus on the virtual model and twin data to enable a thorough monitoring of pumping station energy consumption and carbon emissions. With the help of these models, the operational processes of the pumping
station can be precisely monitored and analyzed with regard to energy consumption and carbon emissions. The digital twin system platform plays a crucial role in providing early warning scheduling for carbon emissions and predicting future outcomes, as demonstrated in Figure 15. This system is an essential tool for enhancing the accuracy of tracking and analyzing energy usage and carbon emissions within the pumping station operations.

Figure 15. The digital twin system platform.

4.5. Analysis and Results

Through an in-depth analysis of this case, we have improved our comprehension of the challenges faced by digital twins in complex infrastructure projects and identified effective strategies to address them.

A major challenge encountered in constructing the digital twin model of a pumping station is the integration of diverse data from multiple sources. The data required for pumping station operation include design data, real-time sensor data, maintenance logs, and other forms of information. These datasets vary in format, time granularity, and semantic interpretation. To address this issue, a seamless integration of the disparate data was achieved through the establishment of data standards and the execution of a unified data preprocessing process.

Another significant challenge lies in accurately replicating the behavioral logic of the pumping station in digital space due to its dynamic and complex process system. To address this challenge, our approach involved a comprehensive utilization of the theory of the digital twin five-dimensional model. This enabled us to establish a real-time mapping between the physical entity and the virtual model by leveraging various data and connection methods. The development of a physical entity-based model enabled us to realistically simulate the operational behavior of the pumping station. Additionally, we utilized data-based connections to analyze historical operational data patterns and forecast potential anomalies. Integration of both these approaches in the digital twin model effectively enables the depiction of the pumping station’s overall operational status with a high level of precision.

Through the operation and tuning of the digital twin model, we have successfully completed the automatic inspection of the robot and conducted a detailed analysis of the optimization space of the pumping station across various dimensions such as energy efficiency, carbon emissions, and operation and maintenance costs. One example of this is the integration of the twin model with the inspection robot simulation path, which facilitated the achievement of 100% automated inspection and resulted in a 75% decrease in inspection time. Consequently, significant cost reductions were achieved through the elimination of the necessity for eight operators, leading to a reduction in operational costs.
of approximately RMB 2.25 million. Furthermore, the implementation of model-based scheduling and optimization facilitates automatic alarms and protection mechanisms, thereby enhancing scheduling stability by around threefold. These tangible economic benefits underscore the practical advantages brought about by the utilization of digital twin technology.

The successful implementation of this case provides valuable insights for advancing high-precision digital twin technology in similar complex infrastructure projects in the future.

(1) Standardization of modeling process: By documenting the standardized modeling process and best practices in this case, it can serve as a guide for future projects, thereby enhancing modeling efficiency.

(2) Migration of modeling methodology: The modeling methodology and key technologies employed in the development of the five-dimensional model can be adapted not only to pumping stations but also to other process-oriented infrastructure sectors such as hydropower stations and petrochemical pipeline networks.

(3) Cross-system and cross-domain model integration: Going forward, there is potential for exploring the integration of digital twin models across different systems and domains to support higher-level decision-making processes.

(4) Integration of emerging technologies: By incorporating emerging technologies such as artificial intelligence, Internet of Things, and 5G, the digital twin model can further improve its applicability, intelligence, and responsiveness.

5. Conclusions

This paper aims to establish a standardized process for constructing a precise and comprehensive digital twin model in the infrastructure field, in order to facilitate the seamless integration of virtual and actual environments. The foundation for this integration lies in the development of the digital twin model that authentically mirrors the real world. Although certain standards and technologies support the creation of such twin models, research in the field of infrastructure-specific digital twin is still relatively nascent. The standardization process proposed in this paper seeks to enhance the precision of these models.

This study has thoroughly examined the process of precise digital twin modeling, leading to the development of a comprehensive set of standards for creating models and a technical framework. Expanding on the foundation established, this study has utilized a particular modeling methodology for an intelligent pumping station project, showcasing accurate model development and offering significant insights. Consequently, this application functions as a valuable resource for further exploring and comprehending the methods and theories of constructing digital twin models in the infrastructure sector. This research provides a practical framework for infrastructure projects intending to deploy digital twin technologies, contributing to the growing body of knowledge in this field. It emphasizes the significance of accurate modeling and standardization in enhancing the capabilities and applications of digital twin technology in infrastructure management and operations. The goal is to connect the virtual and physical realms. Therefore, the results emphasize the significance of these elements in promoting the development of digital twin solutions for infrastructure projects.

Some limitations do exist in this study, which opens up avenues for future research endeavors:

First, despite our efforts to attain high accuracy, the intrinsic complexity and variability of infrastructure systems pose challenges for a single digital twin model to capture all nuances precisely. Future work should explore model fusion techniques, integrating multiple digital twin sub-models with varying resolutions and foci to provide a more comprehensive representation of intricate systems.

Secondly, our current model primarily focuses on static modeling, with limited capabilities in simulating dynamic behaviors. Leveraging emerging technologies such as artificial
intelligence is a promising direction to enhance the model’s simulation and prediction abilities for complex dynamic processes.

Lastly, this work concentrates on digital twin modeling for individual infrastructure systems, lacking extension to broader complex systems such as smart cities and intelligent transportation. Future research should investigate the scalability and application of digital twin models at system-level and global-level scales.

**Author Contributions:** Conceptualization, Z.L.; data curation, F.F. and Y.M.; methodology, F.F.; software, F.F.; validation, FF, Z.L., and G.S.; writing—original draft, F.F.; investigation, Y.M.; resources, Z.L.; writing—review and editing, G.S. and Y.M.; supervision, Z.L.; project administration, Z.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research was funded by the National Key Research and Development Program of China No. 2022YFC3003600.

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

**References**


**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.