Research on Construction Technology of Digital Twin Data Baseboard

Rui Ma [1,2,3,4*], Shuangping Li [1,2,3,4], Chongzheng Yi [1,2,3]

*Corresponding author's e-mail: 58122712@qq.com
Li Shuangping's e-mail: lishuangping@cjwsjy.com.cn
Yi Chongzheng's e-mail: yichongzheng@cjwsjy.com.cn

[1] Changjiang Survey Planning Design And Research Co., Ltd., Wuhan, Hubei, 430010, China
[2]Changjiang Spatial Information Technology Engineering Co., Ltd., Wuhan, Hubei, 430010, China
[3] Hubei Province Engineering Technology Research Center for Water Resources Information Perception and Big Data, Wuhan, Hubei, 430010, China
[4]Hubei Key Laboratory of Basin Water Security, Wuhan, Hubei, 430010, China

Abstract. Accelerating the construction of smart water conservancy is one of the six implementation paths for the high-quality development of water conservancy in the new stage, and also a significant sign of the high-quality development of water conservancy in the new stage. Water conservancy data baseboard is the "calculation data" of intelligent water conservancy construction and the foundation of digital twin water conservancy construction. It is necessary to strengthen the data construction standard, mine the data value, enhance the data service value, consolidate the water conservancy data baseboard, and support the research and development of the "four pre-stages" function of water conservancy business. On the basis of analyzing the requirements of national informatization construction and the objectives of digital twin water conservancy construction of the Ministry of Water Resources, this paper expounds the positioning of the water conservancy digital backboard, puts forward the framework and path of building the water conservancy digital backboard, and carries out relevant research on key technologies, in order to guide data collection, aggregation, management, storage and service, and effectively support the construction of digital twin water conservancy. It provides reference for water conservancy information technology research and digital baseboard construction practice.

Keywords: Smart water conservancy; Digital twin water conservancy; Data backplane; Data governance; Data service

1. Introduction

Since the 18th National Congress of the CPC, China has attached great importance to the development of informatization. In 2016, at the National Network security and Informatization Work Conference, it was proposed to emphasize that we should grasp the historical opportunity of informatization development, promote the construction of a network power through independent innovation, there is no modernization without informatization, and to build a network power, we must have our own technology and excellent technology. China's 14th Five-Year Plan and 2035 Vision Outline emphasize that digital transformation should drive the overall transformation of production, lifestyle and governance, and create new

PMBDA 2023, December 15-17, Nanjing, People's Republic of China

DOI 10.4108/eai.15-12-2023.2345418

advantages of the digital economy. Give full play to the advantages of massive data and rich application scenarios, and promote the deep integration of digital technology and the real economy. Therefore, digital transformation, digital and intelligent construction is very necessary and urgent.

On June 28, 2021, Minister Li Guoying of the Ministry of Water Resources clearly proposed at the "Three pairs of benchmarks and one Plan" special action summary conference that in order to accelerate the construction of smart water conservancy, digitalization, networking and intelligence should be the main line in accordance with the requirements of "demand driving, application first, digital empowerment and capacity enhancement", and digital scenarios, intelligent simulation and precise decision-making should be the path. Comprehensively promote the construction of data, algorithms and computing power, and accelerate the construction of a smart water conservancy system with the functions of "four pre-forecasts" (forecast, early warning, rehearsal and pre-plan)[1]. In order to implement Minister Li Guoying's strategic deployment on the construction of smart water conservancy, The Ministry of Water Resources has issued a series of top-level design, technical guidance and technical requirements, such as the "Top-level design of Smart Water Conservancy Construction", the "14th Five-Year Plan" Implementation Plan for Smart Water Conservancy Construction, the "Technical Outline for Digital Twin Basin Construction", the "Technical Guidelines for Digital Twin Water Conservancy Project Construction", and the "Basic Technical requirements for four pre-construction functions of water conservancy business". These documents also make it clear that the data baseboard is one of the key tasks of the construction of the digital twin basin.

In recent years, with the rapid development and evolution of remote sensing satellites, iot sensing, data mining, data governance and other technologies, there are more and more means of data acquisition, and the update frequency is higher and higher, and data has become an essential and important part of intelligent construction. Data board, data base and other concepts are also emerging in an endless stream[2], such as the "spatial data base" and "three-dimensional data base" of smart city, the "data baseboard" of smart water conservancy construction, the "digital base" of communication industry, and the "data base" proposed by Huawei and other IT technology companies. The construction of tamping data base or data baseboard is of great significance.

Water conservancy data has the characteristics of multi-source (air, sky and earth), multi-scale (multi-level watershed) and multi-dimension (time and space), etc. The construction content involves data collection, data governance, data storage, data management, data display, data application and other links[3-4], and there are many problems to be solved such as spatio-temporal consistency, logical consistency, topological consistency and attribute consistency. A complete theoretical system of data baseboard has not yet been formed[5-8], and the construction of water conservancy data baseboard is faced with new opportunities and challenges. It is necessary to establish a set of unified data baseboard construction standards to strengthen the effectiveness and reliability of data in the application process, effectively manage water conservancy data, comprehensively mine water conservancy data, and enhance the service value of water conservancy data. Support the research and development of the "four pre-stages" function of water conservancy business[7].

2. Construction Ideas

The digital twin water conservancy is based on the physical basin as the unit, the spatiotemporal data as the base, the water conservancy model as the core, and the water conservancy knowledge as the driving force. The digital mapping, intelligent simulation, and multi-scheme optimization of all elements of the physical basin and the whole process of water conservancy management activities are implemented to realize synchronous simulation operation with the physical basin, virtual-real interaction, and iterative optimization, and support accurate decision-making. Data baseboard is the "calculation data" of smart water conservancy construction. Through water conservancy perception network and water conservancy information network, it gathers multi-source massive heterogeneous data sources, provides data services, supports data engine monitoring, and provides a solid "data base" for simulation engine, knowledge engine, and intelligent business application. The positioning and function of water conservancy data baseboard in digital twin water conservancy construction are shown in Figure 1.

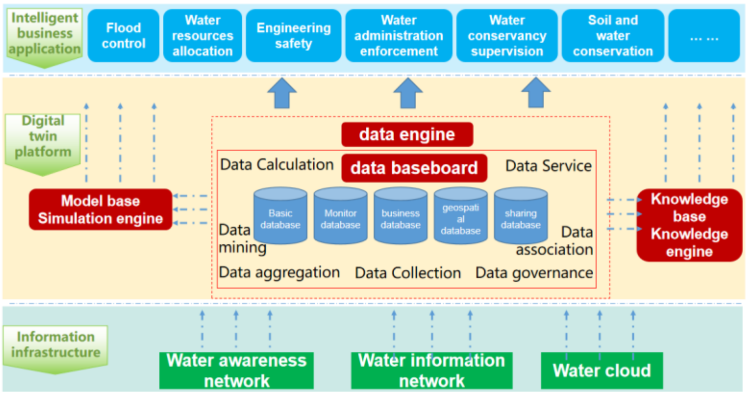

Fig.1 Location of water data baseboard

The construction of water conservancy data baseboard can make full use of the existing data and sharing system, collect water conservancy basic data, business management data, dynamic monitoring data, geospatial data and cross-industry sharing data within the basin or project scope, and form a unified data and two-three-dimensional integrated data baseboard with data model as the core and data engine as the driving force. Realize the digital mapping of all elements, and then integrate with the model platform and knowledge platform to achieve the standardization of business data, forming a data support system with continuous updating ability. The overall architecture of water conservancy data baseboard is shown in Figure 2.

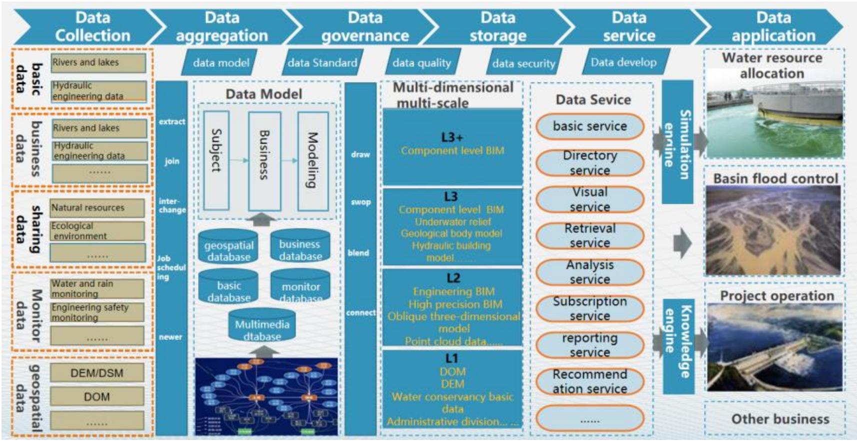

Fig.2 The overall architecture of building water baseboard

3. Construction Path

As can be seen from Figure 2, the construction of water conservancy data baseboard can be realized through steps such as data collection, data aggregation, data storage, data governance, and data service. This series of processes can also be classified as data collection - aggregation - storage - governance - use, so as to build a unified and standard data resource pool, solve the problem of "one-stop" aggregation, retrieval and acquisition of global data resources[9], construct multi-level integrated water conservancy data baseboard.

3.1 Data Collection

Focusing on the construction of water conservancy digital twins or service objects, combined with the data sources collected by the water conservancy perception network, five types of data such as water conservancy basic data, dynamic monitoring data, business management data, geospatial data, and cross-industry sharing data are collected and integrated. The specific contents are as follows:

(1) Water conservancy basic data: including the two-dimensional and three-dimensional data of rivers (river channels, water flows), lakes, reservoirs, embankments, hydrogeology, flood storage areas and other water systems and water conservancy projects within the scope of construction.

(2) Dynamic monitoring data: including water situation, rain situation, industrial situation, water quality, sediment, disaster situation, groundwater level, water access, moisture content, water conservancy project safety operation monitoring data, video, network public opinion and so on.

(3) Business management data: including intelligent business application data and the data generated during its operation.

(4) Geospatial data: including administrative divisions, landforms, land cover, digital orthophoto, digital elevation model, digital surface model, oblique photographic image/laser

point cloud, underwater terrain, building information model (BIM), etc. According to the data accuracy and construction scope is divided into L1, L2, L3 three levels of management.

(5) Cross-industry sharing data: including data sources that need to be shared from the Ministry of Industry and Information Technology, the Ministry of Natural Resources, the Ministry of Ecology and Environment, the Ministry of Housing and Urban-Rural Development, the Ministry of Agriculture and Rural Affairs, the National Bureau of Statistics, the China Meteorological Administration and other relevant departments.

3.2 Data Aggregation

According to the principle of "one data one source", the data convergence standard is clearly defined, and the collected data sources are fully stored into the data lake in the original format. In the process of data entering the lake, it is necessary to determine the aggregation means. Various types of data can be aggregated through data extraction, shared access, network transmission, network download, file collection, and real-time data service. Data aggregation engine can also be developed, and various aggregation means can be encapsulated by workflow mechanism to realize automatic data entering the lake.

(1) Aggregation of basic data and business data. For the data of the established reservoir business (including hydrological information system data, reservoir information management data, etc.) or the shared service platform (reservoir management information release, etc.), the method of "data extraction" can be used to enter the lake.

(2) Monitoring data aggregation. Real-time monitoring data such as rain, video, and work conditions of existing and new monitoring systems (dam monitoring systems, video monitoring systems, etc.) are entered into the lake by "shared access" and the metadata is synchronously entered into the lake.

(3) Geospatial data aggregation. The newly created image map, elevation map, two-three-dimensional model, etc. are collected into the lake in "file form", and the corresponding metadata are input at the same time.

(4) Cross-industry data aggregation. External data such as meteorological data and land use data of socio-economic basin can be imported into the lake by means of "network transmission", "network download", or access to "data service", and the metadata can be imported into the lake simultaneously.

3.3 Data Governance

The data lake formed through data aggregation has problems such as inconsistent standard scales, abnormal data redundancy, and poor quality, and cannot be used directly. Therefore, data governance needs to be carried out to improve data quality, achieve data standardization, consistency, integrity, legitimacy, and availability, and avoid data redundancy and conflicts.

(1) Standardization construction. According to the "SL T 801-2020 Water Conservancy map Spatial Information Service Code", "SL T 701-2021 Water Conservancy Information Classification and Coding General Rules", "SL T 809-2021 Water Conservancy object basic database table structure and identifier" and other norms and standards, the data are

standardized processing. The processing contents include the unification of coordinate system, format, field attribute, service meaning and expression rules.

(2) Data cleaning and transformation. To solve the problems of data errors, duplication, conflict, etc., the cleaning and conversion content includes topological processing, merging and splitting, compressed texture, smoothing, resampling, clipping, spatial connection, integration, etc.

(3) Data fusion processing. According to the data scale, it is divided into L1, L2, L3, L3+ and other levels to realize the spatial superposition and rule association of multi-source (air, sky and ground), multi-scale (multi-level watershed), and multi-dimensional (time and space) data, forming a digital twin space that integrates macro and micro, above-ground, indoor and outdoor, water and underwater, and two and three dimensions.

3.4 Data Storage

After data governance, according to the types and characteristics of data, relational database, spatial database, distributed database and distributed file system are used to store and manage structured, geospatial, semi-structured and unstructured data in a unified manner.

3.5 Data Services

Provide diversified data service capabilities, including resource catalog services, data visualization services, data sharing services, etc.

(1) Resource catalog service: sorting, encoding and describing stored information resources to facilitate retrieval, location and access to information resources.

(2) Data visual services: provide an intuitive and three-dimensional perspective, display data information in an immersive interactive way, so that users can have a glimpse of water information knowledge, so as to effectively assist intelligent management and decision-making, such as three-dimensional map services.

(3) Data sharing service: Using service encapsulation and combination technology to publish data access functions into data services, build a water conservancy data resource service system, and provide users with multi-mode data sharing services such as service aggregation, publication and subscription, and portal access through API interface sharing and front-end sharing.

4. Research on key technologies

The construction of water conservancy data baseboard not only uses a lot of computer science and technology, but also needs multi-disciplinary technology integration and innovation. Through the key technology research, solve the application difficulties and pain points, optimize and accelerate the data baseboard construction speed, improve the data baseboard construction efficiency.

4.1 Deep integration of GIS and BIM

GIS focuses on the data management and display of macro and large scenes, while BIM can better describe the local fine structure. Compared with GIS, BIM has a higher data density within a unit. In order to better support digital twin water conservancy construction, GIS+BIM data integration with macro and micro integration should be realized to build a unified digital scene.

(1) BIM lightweight

On the premise of ensuring that the accuracy of BIM geometric model and texture is not lost and the attribute information is complete[10], a multi-level level of detail (LOD) system in line with the principle of scene proximity is constructed to achieve efficient hierarchical compression of the original BIM model and generate a multi-level LOD 3D model for dynamic scheduling of the simulation engine.

(2) Integration of GIS and BIM scenarios

First of all, a unified geo-spatial reference is established to ensure that BIM model and GIS data are in a set of geo-spatial scenes. Then, GIS data is clipped and filled based on the spatial shape of BIM model to achieve one-click, automatic and seamless integration of BIM and GIS data, thus forming a complete three-dimensional spatial scene[11]. Figure 3 shows the effect of intelligent integration of Catia BIM data and real 3D scenes. In addition, in the scene simulation, combined with 3D space occlusion, instantiation and other technologies[12], users can naturally and smoothly complete the transition from macro to micro when observing the GIS+BIM model, and improve the data loading speed and rendering efficiency.

Fig.3 GIS+BIM Scenario fusion

(3) Integration of GIS and BIM functions

BIM software and GIS software respectively provide rich data management and operation calculation functions, such as BIM calculation, BIM cutting operation, GIS data management, GIS spatial analysis, etc. Compared with BIM software, GIS software provides perfect geospatial reference, has better data management capabilities and secondary development interfaces, and is more in line with the business application requirements of digital twins[13-14]. Therefore, in general, GIS software is selected as the platform for the integration and fusion of GIS+BIM functions, and then relevant BIM functions are accessed through the interface of BIM software, or GIS+BIM functions are targeted to achieve function-level integration.

Figure 4 shows the cutting analysis and component decomposition effects of the BIM model on the GIS+BIM platform.

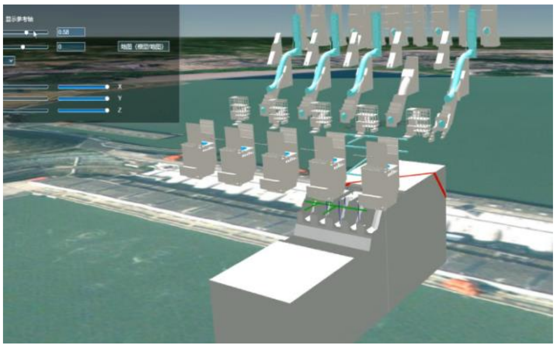

Fig.4 GIS+BIM Function fusion

4.2 Integrated data visualization service

In order to realize synchronous simulation operation and virtual-real interaction with physical water network, river basin and water conservancy project, digital twin water conservancy should provide immersive simulation scenes to meet multi-level and multi-granularity data display and expression. Therefore, after realizing the fusion and management of multi-source massive heterogeneous data, the water conservancy data base should fully consider the resolution scale, information granularity and hierarchical relationship of various data, and establish a technical system for full-space data roaming and visualization[15-16] to meet the requirements of integrated data visualization and twin simulation. The integration of macroscopic and microscopic, above-ground and underground, indoor and outdoor, water surface and underwater, and three-dimensional expression of water conservancy objects in twin scenes is realized, as shown in Figure 5, thus better supporting the multi-dimensional management of water conservancy business.

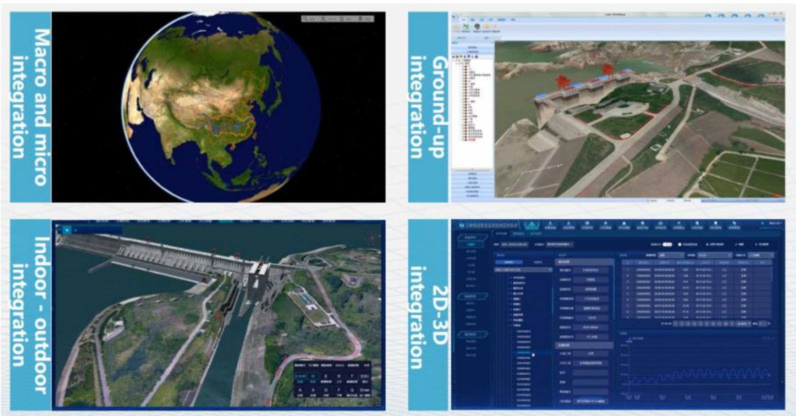

Fig.5 Full spatial visualization

5. Conclusion

Data is the foundation of digitization, informatization and intelligent construction, and the construction of rich, reasonable, solid and reliable data sources and the construction of flexible and scalable data framework system will effectively support the construction of digital twin water conservancy and smart water conservancy. This paper puts forward the construction framework and implementation path of water conservancy digital twin data baseboard, including data collection, data aggregation, data governance, data storage, data service, etc., and carries out related key technology research. The relevant ideas and schemes can provide reference for the data construction and system research and development of digital twin water networks, river basins and water conservancy projects.

With the rapid development and evolution of the new generation of information technology, the deep integration of water science and computer science, geography, environmental science, mathematics and sociology has been promoted, and the iteration of new data governance and service solutions has accelerated. We should embrace changes and adhere to technological innovation. The construction and development ideas of data baseboard with data information as fuel and digital technology as engine are established to further drive water conservancy elements, achieve optimal configuration, the most precise reorganization, and the most efficient operation, and cast the soul for the digital twin water conservancy construction, effectively promoting the development and progress of China's smart water conservancy.

Acknowledgments: This study is supported by The National Key Research and Development Program of China: The key technologies of digital twin platform construction and four-pre-response for small and medium-sized river basins (No.2022YFC3005504)

References

[1] Cai Yang, Cheng Jianguo, et al. alaccelerate the construction of a smart water conservancy system with the function of "four pre-stages"[J]. China Water Resources,2021(20):1-5.

[2] Alif Rachman Harfian, Bagas Satriyotomo, et al. An Investigation of RTOS-Based Sensor Data Management Performance for Tel-USat On Board Data Handling (OBDH) Subsystem[C]. 2019 International Conference on Information and Communications Technology: ICOIACT 2019, Yogyakarta, Indonesia, 24-25 July 2019, [v.2].

[3] Zhou Chao, Tang Haihua, et al. Thinking on the construction of Big Data collection management system in water conservancy industry [J]. Water Conservancy Information Technology, 2021(4): 6-10.

[4] Jiang Yunzhong, Ye Yuntao, et al. Research status and prospect of water conservancy big data [J]. Journal of Hydroelectric Power, 2020, 39(10): 1-32.

[5] Sun H L. Design and implementation of data collection platform in national flood control and drought relief command system [J]. China Flood Control and Drought Relief, 2020, 30(6): 20-26.

[6] Wei Xiangyang, Zhu Jie, et al. Discussion on intelligent flood control in Yellow River Basin [J]. China Flood Control and Drought Relief, 2022, 32(3): 41-46.

[7] CHENG Y L. A preliminary study on the application of big data in water conservancy [J]. Water Conservancy Information Technology, 2019(5): 1-5.

[8] Liu Yes-Sen, Liu Chang-Jun, et al. Construction of digital twin basin data base for "four

pre-forecast" flood control [J]. China Flood Control and Drought Relief, 2022, 32(6): 6-14.
[9] SDN Wessels, R Dixon. Geotechnical data aggregation and visualisation supporting informed risk management: the one-stop geotech shop[C]. Slope Stability 2020: Proceedings of the 2020 International Symposium on Slope Stability in Open Pit Mining and Civil Engineering, 12-14 May 2020, vol.1.
[10] T. W. Kang, C. H. Hong. Development of Lightweight BIM Shape Format Structure to Represent Large Volume Geometry Objects Using GIS with Facility Management[C]. 31st International symposium on automation and robotics in construction and mining: ISARC 2014, Automation, Construction and Environment: Sydney, Australia, 9-11 July 2014.
[11] Zhang Li, Ma Rui, et al. Research and Application of 3D Geographic Information Technology in the construction of Digital Xiaolangdi Project [C]. Proceedings of the 2016 Academic Annual Meeting of the Chinese Hydraulic Society, 2016-10-19: 592-596.
[12] Ma Rui, Qiu Xin, et al. 3DGIS+BIM Technology helps the intelligent application of water conservancy projects throughout their life cycle [J]. China Construction, 2020, 218(7): 73-76.
[13] Ma Rui, Dong Lingyan, et al. Dam safety management platform based on Internet of Things and 3D visualization technology and its implementation [J]. Journal of Yangtze River Scientific Research Institute, 2019, 36(10): 111-116.
[14] Tang Xiangqian, Ma Rui, et al. 3D geographic information platform -- Ark [J]. Journal of Water Resources and Hydropower Letters, 2019, 40(7): 6-7.
[15] XIE Mingxia, LI Li, Ma Rui. Application of 3DGIS+BIM Technology in urban subway Management Information System [J]. Geospatial Information, 2019, 17(9): 86-89,130.
[16] Liu C K, Ma Rui, et al. 3D dynamic prediction of flood risk supported by 3DGIS [J]. Journal of Yangtze River Scientific Research Institute, 2019, 36(10): 117-121.