Towards a Geological Digital Twin: Integrating Multi-Dimensional Geological Models for Comprehensive Earth System Representation

Xuan Wu [1], Fanyu Qi [1], Xiaolei Li [1], Biao Jiang [2], Chao Zhang [1]

[1] Development and Research Center of China Geological Survey, Beijing - wxuan@mail.cgs.gov.cn
[2] Institute of Mineral Resources, Chinese Academy of Geological Sciences, Beijing

Keywords: Geological Digital Twin, Multi-Dimensional Models, Data Integration, Earth System Representation, Spatial-Temporal Data.

Abstract

Understanding and managing Earth's complex geological systems are crucial for sustainable development and disaster mitigation. This paper details the National Geological Data Center's (NGDC) foundational efforts in China towards a comprehensive Geological Digital Twin. Guided by principles of integrated planning, standardization, and technology-driven services, NGDC has built a robust data infrastructure and developed extensive multi-dimensional geological models. We highlight innovative practices in establishing a standardized "receive, manage, utilize" data management system, integrating vast multi-source, multi-scale geological data into cohesive database clusters, and deploying open service platforms, including an internationally recognized geological data publishing system. Key achievements include the construction of China's most extensive 1:50,000 regional geological map database, covering half of the landmass, a comprehensive mineral deposit database, and a widely utilized online geological map service. These integrated multi-dimensional models serve as a digital replica of the subsurface, enhancing understanding, analysis, and decision-making for Earth system representation. We discuss the significant contributions made, the challenges encountered, and future directions for AI-empowered modeling and real-time data assimilation in achieving a truly dynamic and comprehensive Geological Digital Twin.

1. Introduction

The Earth's geological processes are inherently complex, multi-scale, and dynamic, playing a critical role in natural resource distribution, environmental stability, and geological hazards. Accurate, comprehensive, and accessible geological information is thus fundamental for national strategic planning, scientific research, and societal well-being(Kim, 2023). Traditionally, geological data management has often been characterized by fragmentation, heterogeneity, and limited interoperability, hindering a holistic and dynamic understanding of the subsurface. This fragmented landscape makes it challenging to conduct integrated analysis, simulate complex geological processes, or respond rapidly to geological emergencies (Aufaristama, 2025).

The concept of a "Digital Twin" (DT), a virtual replica of a physical system that is continuously updated with real-time data, has emerged as a transformative paradigm in various fields, including manufacturing, urban planning, and infrastructure management(Phelps, 2020). Extending this concept to the geological domain, a "Geological Digital Twin" (GDT) envisions a comprehensive, dynamic, and high-fidelity digital representation of the Earth's geological structures, processes, and resources(Aufaristama, 2025). Such a GDT would enable advanced analysis, simulation, prediction, and intelligent decision-making, moving beyond static maps and databases to a living, evolving digital Earth(Shi, 2022).

This paper outlines the strides made by the National Geological Data Center (NGDC) in China in laying the groundwork for a comprehensive Geological Digital Twin. Guided by principles of integrated planning, standard-driven development, technological innovation, service orientation, and robust security, the NGDC has systematically built a national-level geological data infrastructure. Our work focuses on integrating vast volumes of multi-dimensional geological data, developing sophisticated geological models, and establishing advanced service platforms to facilitate a comprehensive digital representation of the Earth's geological system. This paper details our methodologies, key achievements, the challenges encountered, and our future directions towards realizing a fully functional and dynamic Geological Digital Twin(Liu, 2022).

2. Related Work

The evolution of geological data management has progressed from paper-based archives to digital databases and Geographic Information Systems (GIS). Early efforts focused on digitizing geological maps and creating relational databases for geological attributes. With the advent of advanced computing and sensing technologies, the focus shifted towards integrating multi-source, multi-scale geological data, leading to the development of national geological information systems in many countries(Li, 2023). However, these systems often primarily serve as repositories, lacking the dynamic, real-time, and predictive capabilities inherent in the Digital Twin concept.

The Digital Twin concept originated in manufacturing and has since expanded to smart cities, infrastructure, and even human organs. Its core tenets involve a physical entity, a virtual model, and real-time data connections between them, enabling monitoring, simulation, and optimization. Applying the DT concept to the geological domain presents unique challenges due to the inherent complexity, inaccessibility, and vastness of the subsurface. Geological data is often sparse, uncertain, and multi-dimensional (e.g., 2D maps, 3D models, 4D processes). Existing research on geological modeling often focuses on specific aspects like 3D geological modeling for resource estimation or mineralization model simulation, but rarely encompasses a holistic, dynamic, and continuously updated "twin" of the entire geological system at a national scale(Jiang, 2022).

The NGDC's work distinguishes itself by its national scope, comprehensive data integration across diverse geological disciplines and scales, and the establishment of a professionally recognized data publishing platform. These efforts directly address the foundational requirements for building a Geological Digital Twin, moving beyond mere data storage to active data management, integration, and service provision that can support dynamic Earth system representation.

3. Methodology and Approach

The construction of the NGDC's data infrastructure and the foundation for a Geological Digital Twin is underpinned by a systematic approach encompassing data lifecycle management, multi-dimensional data integration, and advanced service platform development(Kim, 2023; Zhu, 2021). This holistic strategy ensures data quality, accessibility, and utility for diverse applications(Wright, 2020).

3.1 Normative "Receive, Manage, Utilize" Data Management System

A cornerstone of our approach is the establishment of a standardized and highly efficient data management system that governs the entire data lifecycle. This system is meticulously designed to ensure data integrity, traceability, and security from the moment of collection to its ultimate utilization.

3.1.1 Strict Adherence to National Regulations and Standards: We strictly adhere to the "Geological Data Management Regulations" and a comprehensive suite of national and industry standards for geological data submission. This includes adherence to data formats , coordinate systems, and thematic classification schemes. This rigorous compliance ensures the consistency and interoperability of data collected from various geological work projects across the country, forming a unified national dataset.

3.1.2 Meticulous Data Validation: Each submitted dataset undergoes a multi-layered validation process. Includes: Technical Checks: Verifying data format compliance, projection accuracy, and file integrity. Logical Checks: Ensuring internal consistency within the dataset (e.g., topological correctness of polygons, connectivity of lines, attribute domain validity). Scientific Review: Expert geologists review the data for scientific accuracy and consistency with known geological principles and regional geological settings. This comprehensive quality control process is critical for building reliable multi-dimensional models.

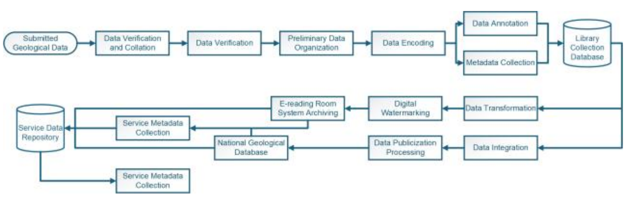

Figure 1. Data Management and Utilization Workflow.

3.1.3 Comprehensive Metadata Management: Beyond basic data, comprehensive metadata is created or meticulously supplemented for every dataset. This metadata includes information on data origin, acquisition methods, processing history, quality indicators, spatial and temporal extent, and thematic content. Standardized metadata schemas (e.g., ISO 19115, FGDC) are employed to ensure discoverability, understanding, and long-term preservation of the data, which is vital for the reusability and interpretability of complex multi-dimensional models within a GDT.

3.1.4 Standardized Archiving and Traceable Management: Qualified data is then ingested into a standardized national geological database, utilizing robust spatial database technologies. Data is organized according to unified data schemas and storage protocols, ensuring consistency and efficient retrieval. This structured archiving enables full traceability, allowing users to track the origin, processing lineage, and quality history of any dataset, which is fundamental for the trustworthiness and auditability of a digital twin.

3.1.5 Data Security and Intellectual Property Protection: To safeguard sensitive geological information while promoting responsible sharing, multi-layered security measures are implemented. This includes access control mechanisms based on user roles and data sensitivity levels, encryption for data in transit and at rest, and the embedding of hidden watermarks into service data. These watermarks serve as a deterrent against unauthorized use and provide a means to track data dissemination, thereby protecting national intellectual property rights and balancing the need for open science with national security imperatives.

3.2 Large-Scale Integration and Multi-Dimensional Data Product Development

The core of building multi-dimensional geological models lies in the systematic integration of diverse, multi-source, and multi-scale geological data. This process transforms raw, disparate datasets into cohesive, interoperable information products that form the building blocks of the Geological Digital Twin.

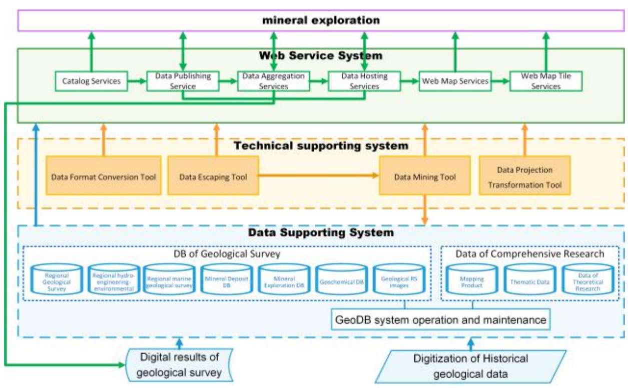

Figure 2. Big data service framework.

https://doi.org/10.5194/isprs-archives-XLVIII-4-W14-2025-327-2025 | © Author(s) 2025. CC BY 4.0 License.

3.2.1 National Basic Geological Database Clusters: We systematically extract, organize, and integrate core thematic data, including basic geology (e.g., lithology, stratigraphy, geological structures), geochemistry, hydrogeology, geological hazards, and mineral deposits. This integrated data is continuously updated into the National Geological Database. We have constructed extensive database clusters covering medium to large scales (1:50,000, 1:200,000, 1:250,000) for regional geological surveys, mineral surveys, hydrogeology, engineering geology, environmental geology, and geochemistry. Notably, the 1:50,000 regional geological map database now covers over half of China's terrestrial area, making it the most extensive of its kind domestically. These databases form the fundamental multi-dimensional models representing the Earth's surface and shallow subsurface, providing spatial, attribute, and temporal dimensions.

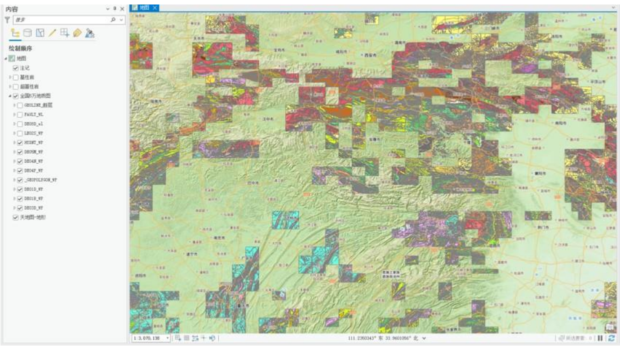

Figure 3. Regional geological map database.

3.2.2 Deep Development of Thematic Databases: Beyond basic integration, we engage in deep development of specialized thematic databases using advanced data mining and association techniques. For instance, we have built the nation's largest and most comprehensive mineral deposit database, containing over 370,000 records. This rich dataset has enabled the production of national resource distribution atlases for 45 mineral types and their online public information services, providing critical insights into resource potential. Furthermore, we conducted large-scale profile data mining and collation work to explore the establishment of a geological profile data model. To date, this model facilitates the integration and management of over 100,000 profile datasets, providing crucial vertical dimension information for subsurface modeling. We also completed the national geological survey work intensity database and maps, covering 7 types of thematic work, efficiently supporting higher-level business departments' statistics and information reporting.

Figure 4. Mineral deposit database.

3.3 Open National Geological Data Center Service Platform

To ensure the accessibility, usability, and impact of these multi-dimensional models, we have developed an open service platform. This platform acts as the interface between the complex data infrastructure and diverse user communities.

3.3.1 Data Center Portal: Launched in April 2021 (http://dc.ngac.org.cn), the NGDC portal serves as a unified gateway for geological data products and services. It incorporates stringent online data release information security mechanisms to protect sensitive data while facilitating access. The portal currently offers a wide array of services, including over 4,700 network digital map services, online browsing of over 200,000 library collections, and access to over 10 database applications, alongside rapid emergency response services.

Figure 5. Data center portal.

3.3.2 International Geological Data Publishing Platform: A pioneering achievement is the establishment of a geological scientific data publishing system. This system integrates data submission, rigorous security review, quality peer review, formal publication, secure storage, citation, and evaluation. It leverages DOI (Digital Object Identifier) to ensure persistent identification and citation of China's geological digital resources, safeguarding intellectual property. To date, 9 data issues have been successfully published, releasing 154 datasets and openly sharing 265 geological survey reports. This platform has achieved significant international recognition, becoming a formal member of the International Science Council World Data System (ISC-WDS) and being indexed by Clarivate Analytics' Data Citation Index (DCI). This significantly enhances the international influence and visibility of China's geological scientific data, fostering global scientific collaboration.

Figure 6. Geological data publishing platform.

3.3.3 Innovative Online Internet Geological Map Data Services: Since 2013, we have pioneered online map services in China by providing medium to large-scale geological map services via OGC-compliant WMS/WMTS protocols. These services are widely invoked by various applications and have become an essential base map for field geologists, demonstrating the practical utility and interoperability of our multi-dimensional spatial data.

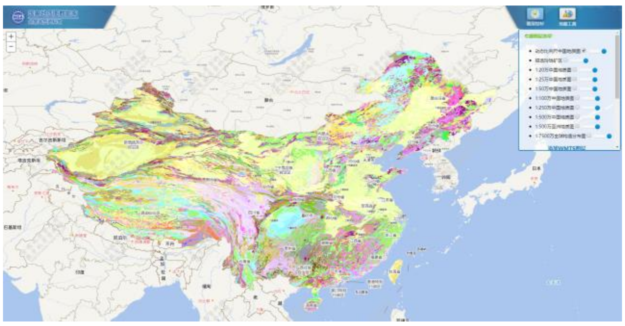

Figure 7. Online internet geological map data services.

4. Results and Key Contributions

The NGDC's comprehensive efforts have laid a robust foundation for a Geological Digital Twin, making significant contributions to Earth system representation.

Comprehensive Multi-Dimensional Data Coverage: The integrated database clusters, particularly the 1:50,000 regional geological map database, provide an unprecedented spatial coverage and thematic depth. This forms the core digital representation of China's geological landscape, enabling a holistic view of geological features and processes.

Rich Thematic Models: The specialized mineral deposit and profile databases offer detailed, attribute-rich multi-dimensional models. These are crucial for precise resource assessment, subsurface characterization, and understanding geological evolution, providing the "intelligence" layer for the digital twin(Xiong, 2023).

Enhanced Data Accessibility: The advanced service platform, including the portal and WMS/WMTS services, ensures that these complex multi-dimensional models are readily accessible and interoperable for various applications, from scientific research to practical engineering.

International Recognition: The geological data publishing platform's international standing significantly promotes the global sharing and citation of China's geological data. This not only enhances our scientific influence but also contributes to a broader, collaborative Earth system understanding.

Practical Application: The integrated geological big data has directly supported national strategic mineral exploration efforts. By leveraging big data analysis, we have enabled rapid screening and evaluation of prospective blocks, demonstrating the tangible utility and economic value of these multi-dimensional models in real-world applications.

5. Discussion of Challenges

Despite these significant achievements, realizing a fully dynamic and comprehensive Geological Digital Twin presents ongoing challenges that require continuous innovation and development(Gong, 2024).

Data Heterogeneity and Semantic Interoperability: The inherent diversity of geological data, originating from various disciplines, acquisition methods, and historical contexts, leads to significant heterogeneity in formats, scales, and semantics(Su, 2023). For instance, integrating legacy paper maps with modern digital surveys, or reconciling different lithological classification schemes, remains a complex task. Achieving true semantic interoperability across these disparate datasets, potentially through the development of geological ontologies or knowledge graphs, is crucial for seamless integration into a unified and intelligent digital twin.

Dynamic Updates and Real-time Assimilation: The subsurface is not static; it undergoes continuous geological processes (e.g., active fault movements, groundwater level fluctuations, land deformation due to subsidence or uplift, seismic activity, volcanic eruptions, landslide movements). Capturing and integrating these real-time geological processes into the digital twin requires advanced sensing technologies (e.g., satellite interferometry, in-situ IoT sensors, distributed fiber optic sensing), robust data transmission infrastructure, and sophisticated data assimilation techniques to update the digital model dynamically. This is a major hurdle for achieving a truly "living" digital twin(Bravo-Haro, 2022).

Computational Demands: Building, rendering, and simulating high-resolution 3D/4D geological models for vast geographical areas demands immense computational power, efficient data structures (e.g., voxel models, tetrahedral meshes), and advanced visualization techniques(Zhai, 2024). The sheer volume and complexity of multi-dimensional geological data, especially when considering temporal changes (4D), pose a substantial computational challenge that often exceeds conventional computing capabilities. Solutions may involve leveraging cloud computing, high-performance computing (HPC), and distributed computing architectures.

Data Governance and Security: Balancing the imperative for open data sharing and scientific collaboration with national security concerns and intellectual property protection for sensitive geological information (e.g., strategic mineral reserves, critical infrastructure sites, geological stability of sensitive areas) is a continuous and evolving challenge. Robust data governance frameworks, granular access control policies, advanced cybersecurity measures, and legal frameworks for data ownership and usage are essential to manage these complexities effectively.

6. Future Work and Outlook

Future work at NGDC will focus on addressing these challenges and advancing the Geological Digital Twin concept through the following measures.

Enhanced AI Integration: Leveraging advanced AI and machine learning techniques for more autonomous and intelligent analysis within the digital twin(Huynh, 2021). This includes deep learning for automated geological feature recognition from imagery and maps, natural language processing for extracting structured information from unstructured geological reports, and reinforcement learning for optimizing exploration strategies or disaster response plans.

Real-time Data Assimilation: Developing sophisticated methodologies and systems for integrating real-time sensor data and monitoring information (e.g., from GNSS stations, seismic networks, drone surveys, satellite remote sensing) to enable dynamic updates and predictive simulations within the Geological Digital Twin. This will allow for real-time forecasting of geological events and changes(Lyu, 2024).

Higher Resolution and Deeper Subsurface Modeling: Advancing techniques for constructing more detailed and deeper 3D/4D geological models, particularly for complex geological structures, deep-seated resources (e.g., geothermal, deep mineral deposits), and urban underground spaces. This involves integrating advanced geological modeling software, implicit modeling techniques, and geostatistical methods(Li, 2022).

Broader Earth System Integration: Expanding the Geological Digital Twin to integrate seamlessly with other domain-specific digital twins, such as atmospheric, oceanic, hydrological, and biological digital twins(Meentemeyer, 2023). This interdisciplinary integration is key to contributing to a holistic Earth System Digital Twin, enabling comprehensive environmental and resource management, and addressing global challenges like climate change and sustainable development.

Innovative Applications: Exploring novel and impactful applications of the Geological Digital Twin in various sectors. This includes supporting smart mining operations, optimizing urban underground space planning, assessing geothermal energy potential, selecting and monitoring carbon capture and storage sites, and enhancing disaster resilience through more precise risk assessment and emergency response. Furthermore, fostering greater user engagement through intuitive visualization and interactive tools will be crucial(Zhang, 2024).

7. Conclusion

The National Geological Data Center has established a strong and pioneering foundation for a Geological Digital Twin through its comprehensive multi-dimensional geological data integration, standardized management, and advanced service platform development. By building extensive and comprehensive geological database clusters and fostering international data sharing, we are creating a high-fidelity digital replica of China's geological system. This work significantly enhances our ability to understand, analyze, and manage the complex subsurface. While challenges remain in achieving dynamism and real-time capabilities, our ongoing efforts in AI integration, real-time data assimilation, and advanced modeling are paving the way for a more intelligent, predictive, and comprehensive Earth system representation. This will ultimately support sustainable development, enhance disaster resilience, and drive scientific advancement in the geological domain.

References

Aufaristama, M. (2025). Use Case of Non-Fungible Tokens (NFTs): A Blockchain Approach for Geological Data Dissemination. *J*. https://doi.org/10.3390/j8010001.

Bravo-Haro, M., Broo, D., & Schooling, J. (2022). Design and implementation of a smart infrastructure digital twin. *Automation in Construction*. https://doi.org/10.1016/j.autcon.2022.104171.

Gong, H., Chen, X., Zeng, S., & Su, D. (2024). Advancements in digital twin modeling for underground spaces and lightweight geometric modeling technologies. *Automation in Construction*. https://doi.org/10.1016/j.autcon.2024.105578.

Huynh, K., Svenning, J., Maitner, B., Lien, A., Kraft, N., Boyle, B., Violle, C., Sandel, B., Neves, D., Burger, J., Weitemier, T., López-Hoffman, L., Hannah, L., Serra-Diaz, J., Núñez-Regueiro, M., Jørgensen, P., Li, Y., Lovett, J., Merow, C., Ernst, K., Baldwin, E., Newman, E., Wiser, S., Breshears, D., Marquet, P., Peet, R., Foden, W., Gallagher, R., McGill, B., Feng, X., Oliveira-Filho, A., Pillet, M., Enquist, B., Roehrdanz, P., Š í mová, I., Park, D., & Morueta‐ Holme, N. (2021). A review of the heterogeneous landscape of biodiversity databases: opportunities and challenges for a synthesized biodiversity knowledge base. *Global Ecology and Biogeography*. https://doi.org/10.32942/osf.io/9zt4a.

Jiang, H., Wei, F., & Sun, J. (2022). Design Path and Application Analysis of Integrated Geological Data Management Platform. *Academic Journal of Environment & Earth Science*. https://doi.org/10.25236/ajee.2022.040510.

Kim, I., Khan, M., & Seo, J. (2023). A boundary and voxel-based 3D geological data management system leveraging BIM and GIS. *Int. J. Appl. Earth Obs. Geoinformation*, 118, 103277. https://doi.org/10.1016/j.jag.2023.103277.

Kim, S., & Ahn, E. (2023). Digital Twin Application and Bibliometric Analysis for Digitization and Intelligence Studies in Geology and Deep Underground Research Areas. *Data*, 8, 73. https://doi.org/10.3390/data8040073.

Li, C., Jessell, M., Wu, L., Wang, Z., Wang, J., Guo, J., & Li, F. (2022). Multiple-Point Geostatistics-Based Three-Dimensional Automatic Geological Modeling and Uncertainty Analysis for Borehole Data. *Natural Resources Research*, 31, 2347 - 2367. https://doi.org/10.1007/s11053-022-10071-6.

Li, Z., Jiang, M., Mao, J., Shu, X., Wang, Y., Wang, X., & Tong, Y. (2023). Integrated Management Technology of Geophysics and Geology Data. *Advances in Engineering Technology Research*. https://doi.org/10.56028/aetr.7.1.158.2023.

Liu, M., Pirasteh, S., Guo, Y., Wang, W., Wang, Q., Zhou, R., Zhu, Q., Zheng, W., Wu, H., Ding, Y., & Zhang, L. (2022). Multi-level voxel representations for digital twin models of tunnel geological environment. *Int. J. Appl. Earth Obs. Geoinformation*, 112, 102887. https://doi.org/10.1016/j.jag.2022.102887.

Lyu, B., Shi, C., & Wang, Y. (2024). Multi-scale generative adversarial networks (GAN) for generation of three-dimensional subsurface geological models from limited boreholes and prior geological knowledge. *Computers and Geotechnics*. https://doi.org/10.1016/j.compgeo.2024.106336.

Meentemeyer, R., Tateosian, L., Vukomanovic, J., White, C., Petrasova, A., Petras, V., & Mitásová, H. (2023). An open-source platform for geospatial participatory modeling in the cloud. Environ. Model. *Softw*., 167, 105767. https://doi.org/10.1016/j.envsoft.2023.105767.

Phelps, G. (2020). NO ONE ANSWER: CHALLENGES TO CREATING A GEOLOGICAL DIGITAL TWIN. . https://doi.org/10.1130/abs/2020am-358337.

Senechal, D., Patricot, C., Pipeau, Y., Cochet, S., Stutz, C., Hourez, B., Delvaux, R., Fournier, Y., Calgaro, B., Koren, C., Lafaurie, C., Olita, P., Vaglio-Gaudard, C., Lenain, R., Vezzoni, B., Caillaux, A., Farda, A., & Coster, S. (2024). The Digital Nuclear Reactor Initiative: First steps towards advanced applications. *EPJ Web of Conferences*. https://doi.org/10.1051/epjconf/202430205009.

Shi, C., & Wang, Y. (2022). Data-driven construction of Three-dimensional subsurface geological models from limited Site-specific boreholes and prior geological knowledge for underground digital twin. *Tunnelling and Underground Space Technology*. https://doi.org/10.1016/j.tust.2022.104493.

Su, Y., Guo, H., Huang, C., Xiao, Q., Li, X., Ran, Y., Yuan, S., Su, J., Feng, M., Liu, F., & Shen, H. (2023). Big Data in Earth system science and progress towards a digital twin. *Nature Reviews Earth & Environment*, 4, 319 - 332. https://doi.org/10.1038/s43017-023-00409-w.

Wright, L., & Davidson, S. (2020). How to tell the difference between a model and a digital twin. *Advanced Modeling and Simulation in Engineering Sciences*, 7. https://doi.org/10.1186/s40323-020-00147-4.

Xiong, Y., Wang, J., Wang, Z., Kreuzer, O., & Zuo, R. (2023). A New Generation of Artificial Intelligence Algorithms for Mineral Prospectivity Mapping. *Natural Resources Research*, 32, 1859 - 1869. https://doi.org/10.1007/s11053-023-10237-w.

Zhai, Y., Wei, S., Li, W., Hou, J., Ji, D., & Cheng, Z. (2024). Construction and application of multi-scale 3D geological models: an example from Qingdao City, China. *Frontiers in Earth Science*. https://doi.org/10.3389/feart.2024.1435824.

Zhang, G., Yu, K., Young, 0., Huang, C., Liu, Y., Wang, Y., Yue, S., Gu, Z., Zhang, G., Ren, X., Liu, P., Liu, Q., Xie, W., Chang, J., Hu, X., Yu, T., Xu, Y., Zhu, J., Li, C., Yang, S., Liu, Z., Li, H., Chen, J., Nie, P., Cai, Y., Niu, X., Huang, W., Yang, S., Dai, Z., & Chen, B. (2024). Yi: Open Foundation Models by 01.AI. *ArXiv*, abs/2403.04652. https://doi.org/10.48550/arXiv.2403.04652.

Zhu, H., Wang, J., Li, X., & Li, P. (2021). Complex 3D geological modeling based on digital twin. *IOP Conference Series: Earth and Environmental Science*, 861. https://doi.org/10.1088/1755-1315/861/7/072046.