REVIEW

Artificial intelligence for geoscience: Progress, challenges, and perspectives

Tianjie Zhao,[1,30] Sheng Wang,[2,30] Chaojun Ouyang,[3,27,30] Min Chen,[4,30] Chenying Liu,[5,30] Jin Zhang,[6,30] Long Yu,[2,30] Fei Wang,[7,27,30] Yong Xie,[8,30] Jun Li,[2,30] Fang Wang,[9,27,28] Sabine Grunwald,[10] Bryan M. Wong,[11] Fan Zhang,[12] Zhen Qian,[4] Yongjun Xu,[7,27] Chengqing Yu,[7,27] Wei Han,[2] Tao Sun,[7] Zezhi Shao,[7,27] Tangwen Qian,[7,27] Zhao Chen,[7] Jiangyuan Zeng,[1] Huai Zhang,[13] Husi Letu,[1] Bing Zhang,[1] Li Wang,[1] Lei Luo,[14] Chong Shi,[1] Hongjun Su,[15] Hongsheng Zhang,[16] Shuai Yin,[1] Ni Huang,[1] Wei Zhao,[1] Nan Li,[17,18] Chaolei Zheng,[1] Yang Zhou,[19] Changping Huang,[1] Defeng Feng,[27] Qingsong Xu,[5] Yan Wu,[20,27] Danfeng Hong,[1,27] Zhenyu Wang,[21] Yinyi Lin,[16] Tangtang Zhang,[22] Prashant Kumar,[25,26] Antonio Plaza,[23] Jocelyn Chanussot,[24] Jiabao Zhang,[9,27] Jiancheng Shi,[29] and Lizhe Wang[2,*]

*Correspondence: lizhe.wang@gmail.com

Received: January 15, 2024; Accepted: August 17, 2024; Published Online: August 22, 2024; https://doi.org/10.1016/j.xinn.2024.100691

GRAPHICAL ABSTRACT

PUBLIC SUMMARY

- What does AI bring to geoscience? AI has been accelerating and deepening our understanding of Earth Systems in an unprecedented way, including the atmosphere, lithosphere, hydrosphere, cryosphere, biosphere, anthroposphere and the interactions between spheres.
- What are the noteworthy challenges of AI in geoscience? As we embrace the huge potential of AI in geoscience, several challenges arise including reliability and interpretability, ethical issues, data security, and high demand and cost.
- What is the future of AI in geoscience? The synergy between traditional principles and modern AI-driven techniques holds immense promise and will shape the trajectory of geoscience in upcoming years.

Artificial intelligence for geoscience: Progress, challenges, and perspectives

Tianjie Zhao,[1,30] Sheng Wang,[2,30] Chaojun Ouyang,[3,27,30] Min Chen,[4,30] Chenying Liu,[5,30] Jin Zhang,[6,30] Long Yu,[2,30] Fei Wang,[7,27,30] Yong Xie,[8,30] Jun Li,[2,30] Fang Wang,[9,27,28] Sabine Grunwald,[10] Bryan M. Wong,[11] Fan Zhang,[12] Zhen Qian,[4] Yongjun Xu,[7,27] Chengqing Yu,[7,27] Wei Han,[2] Tao Sun,[7] Zezhi Shao,[7,27] Tangwen Qian,[7,27] Zhao Chen,[7] Jiangyuan Zeng,[1] Huai Zhang,[13] Husi Letu,[1] Bing Zhang,[1] Li Wang,[1] Lei Luo,[14] Chong Shi,[1] Hongjun Su,[15] Hongsheng Zhang,[16] Shuai Yin,[1] Ni Huang,[1] Wei Zhao,[1] Nan Li,[17,18] Chaolei Zheng,[1] Yang Zhou,[19] Changping Huang,[1] Defeng Feng,[27] Qingsong Xu,[5] Yan Wu,[20,27] Danfeng Hong,[1,27] Zhenyu Wang,[21] Yinyi Lin,[16] Tangtang Zhang,[22] Prashant Kumar,[25,26] Antonio Plaza,[23] Jocelyn Chanussot,[24] Jiabao Zhang,[9,27] Jiancheng Shi,[29] and Lizhe Wang[2,*]

[1]Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
[2]School of Computer Science, China University of Geosciences, Wuhan 430078, China
[3]State Key Laboratory of Mountain Hazards and Engineering Resilience, Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610299, China
[4]Key Laboratory of Virtual Geographic Environment (Ministry of Education of PRC), Nanjing Normal University, Nanjing 210023, China
[5]Data Science in Earth Observation, Technical University of Munich, 80333 Munich, Germany
[6]The National Key Laboratory of Water Disaster Prevention, Yangtze Institute for Conservation and Development, Hohai University, Nanjing 210098, China
[7]Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
[8]School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China
[9]State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China
[10]Soil, Water and Ecosystem Sciences Department, University of Florida, PO Box 110290, Gainesville, FL, USA
[11]Materials Science Engineering Program Cooperating Faculty Member in the Department of Chemistry and Department of Physics Astronomy, University of California, California, Riverside, CA 92521, USA
[12]Institute of Remote Sensing and Geographical Information System, School of Earth and Space Sciences, Peking University, Beijing 100871, China
[13]Key Laboratory of Computational Geodynamics, University of Chinese Academy of Sciences, Beijing 100049, China
[14]International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China
[15]College of Geography and Remote Sensing, Hohai University, Nanjing 211100, China
[16]Department of Geography, The University of Hong Kong, Hong Kong 999077, SAR, China
[17]Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control, Nanjing 210044, China
[18]School of Environmental Science and Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China
[19]Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Key Laboratory of Meteorological Disaster, Ministry of Education, Nanjing University of Information Science and Technology, Nanjing 210044, China
[20]Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China
[21]Department of Catchment Hydrology, Helmholtz Centre for Environmental Research – UFZ, Halle (Saale) 06108, Germany
[22]Key Laboratory of Land Surface Process and Climate Change in Cold and Arid Regions, Chinese Academy of Sciences, Lanzhou 730000, China
[23]Hyperspectral Computing Laboratory, University of Extremadura, 10003 Caceres, Spain
[24]University Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France
[25]Global Centre for Clean Air Research (GCARE), School of Sustainability, Civil and Environmental Engineering, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford GU2 7XH, UK
[26]Institute for Sustainability, University of Surrey, Guildford GU2 7XH, Surrey, UK
[27]University of Chinese Academy of Sciences, Beijing 100049, China
[28]Department of Chemistry, Technical University of Munich, 85748 Munich, Germany
[29]National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China
[30]These authors contributed equally
*Correspondence: lizhe.wang@gmail.com

Received: January 15, 2024; Accepted: August 17, 2024; Published Online: August 22, 2024; https://doi.org/10.1016/j.xinn.2024.100691

Citation: Zhao T., Wang S., Ouyang C., et al., (2024). Artificial intelligence for geoscience: Progress, challenges, and perspectives. The Innovation **5(5)**, 100691.

This paper explores the evolution of geoscientific inquiry, tracing the progression from traditional physics-based models to modern data-driven approaches facilitated by significant advancements in artificial intelligence (AI) and data collection techniques. Traditional models, which are grounded in physical and numerical frameworks, provide robust explanations by explicitly reconstructing underlying physical processes. However, their limitations in comprehensively capturing Earth's complexities and uncertainties pose challenges in optimization and real-world applicability. In contrast, contemporary data-driven models, particularly those utilizing machine learning (ML) and deep learning (DL), leverage extensive geoscience data to glean insights without requiring exhaustive theoretical knowledge. ML techniques have shown promise in addressing Earth science-related questions. Nevertheless, challenges such as data scarcity, computational demands, data privacy concerns, and the "black-box" nature of AI models hinder their seamless integration into geoscience. The integration of physics-based and data-driven methodologies into hybrid models presents an alternative paradigm. These models, which incorporate domain knowledge to guide AI methodologies, demonstrate enhanced efficiency and performance with reduced training data requirements. This review provides a comprehensive overview of geoscientific research paradigms, emphasizing untapped opportunities at the intersection of advanced AI techniques and geoscience. It examines major methodologies, showcases advances in large-scale models, and discusses the challenges and prospects that will shape the future landscape of AI in geoscience. The paper outlines a dynamic field ripe with possibilities, poised to unlock new understandings of Earth's complexities and further advance geoscience exploration.

INTRODUCTION

Geoscientists tackle the most significant environmental, scientific, and societal challenges related to Earth.[1,2] Despite extensive research, several questions remain unanswered, such as the origin of Earth and life[3,4] or the snowball/faint sun paradox,[5] among others.[6–9] Unraveling these mysteries requires modeling a complex geosystem,[10] where Earth presents complicated spatial patterns shaped by diverse interacting processes, including natural sub-systems (such as the biosphere, atmosphere, and lithosphere) and various human activities.[11–13] For instance, predicting geohazards necessitates considering not only the inherent complexities of the geosystem but also the significant influence of activities across multiple spatial scales.[14] Moreover, the geosystem is an ever-evolving network characterized by non-linear processes of high dynamical instability,[15] where inherently stochastic features impose significant constraints on temporal analyses.[16] In this context, while current weather forecasting can achieve relatively accurate predictions over several days, the challenge of making reliable predictions intensifies when extending the time frame to months or longer.[17] Throughout history, geoscience has undergone a transition from the reliance on physics-based models to the utilization of data-driven machine learning (ML) approaches when tackling these challenges. This shift has been facilitated by remarkable advancements in artificial intelligence (AI),[18] data collection techniques,[19–21] and computing resources.[22]

Physics-based models

Geoscientific research fundamentally relies on conceptual models that describe key processes and their interactions,[23] which are subsequently tested using physical and numerical models.[24] Physical models simulate environmental conditions in a laboratory setting,[25] allowing researchers to manipulate variables in a controlled manner and investigate hypothetical scenarios within a large-scale and complicated geosystem.[24] While physical models are effective in certain cases (e.g., using clay models to verify the orogenic theory[26]), they also encounter discrepancies between the controlled virtual laboratory environment and real-world situations.[27] Numerical models condense natural processes into mathematical representations.[28] These equations, designed to mirror the intricate characteristics of the real geosystem, are too complex to be solved analytically[29] and are generally tackled by numerical simulations, such as numerical weather prediction models.[30] Traditional physics-based models aim to uncover hidden mechanisms by reconstructing physical processes, and can provide robust explanations once successfully founded. However, the inherent complexity of the geosystem, coupled with our limited understanding, poses a significant challenge.[31] Making comprehensive assumptions about related factors and their dependencies thus becomes difficult.[32,33] The sophistication and uncertainty in optimizing such models greatly hinders their practical application.[34]

Data-driven approaches

As the availability of geoscience data continues to expand, modern geoscientific challenges are increasingly centered around managing extensive datasets, often with limited or no underlying theoretical knowledge.[17,35,36] In this context, AI demonstrates significant potential.[22,37,38] ML, as a major subfield of AI, is deeply rooted in applied statistics and constructs computational models based on inference and pattern recognition rather than physical rules.[39,40] Typical examples include Gaussian-process-based "Kriging" interpolation,[41–43] the utilization of support vector machines for identifying geomorphological features,[44] and so on.[45–47] The success of these methods has sparked broad interest among geoscientists in employing ML to address Earth science challenges, allowing them to bypass the explicit modeling of physical processes.[44,48,49] While conventional ML methods can effectively handle small-scale problems, they often encounter limitations in more complicated scenarios, particularly when dealing with large volumes of data and broader scales.[50] In this case, deep learning (DL) has brought significant advances[51–53] since AlexNet decisively won the ImageNet challenge in 2012.[54] Beyond applications of convolutional neural networks and Vision Transformers (ViTs),[55,56] densely/fully connected networks have proven useful in tasks such as soil mapping,[57] while recurrent neural networks, including long short-term memory (LSTM) networks, are particularly well suited for time series data and temporal problems.[58] AI models hold promise for advancing modern geoscientific research by learning hidden features directly from data without requiring comprehensive physical prior knowledge. DL, as the primary data mining tool in the big data era, propels the application of AI to geoscience. However, AI techniques still face several challenges, including the notorious data-hungry characteristics, the increased demand for computational resources, and the inherent black-box nature of AI algorithms.[59,60] Addressing these challenges is crucial to further explore the potential of AI in geoscience.

Advanced AI techniques

Solely relying on either physics-based or data-driven models proves insufficient for knowledge discovery in geoscience.[61] Hybrid models or physics-guided/informed/aware ML offer a promising solution by integrating domain knowledge to refine AI models in geoscience.[62] These models incorporate constraints derived from domain-specific insights, such as encoding differential equations from data[63] or imposing physical constraints on data-driven models.[64,65] This integration allows for performance comparable with pure data-driven approaches but with the advantage of requiring less training data.[66] Despite their potential to bridge interdisciplinary gaps between data-driven and physics-based models, the effective implementation of hybrid models remains an open question.[11,67] In addition, the recent success of ChatGPT has emphasized the potential of foundation models to enhance a wide range of tasks.[68] The vast expansion of data in geoscience provides a solid groundwork for the emergence of large geoscientific models.[69] These large models offer new avenues for extracting new insights from data to enrich our understanding of the Earth. Nevertheless, their development is still in the early stage.[70,71] Geodata possesses unique characteristics, such as geo-references, various attribute features, and temporal constraints, which make it challenging to directly apply prominent language- and image-processing techniques from other fields to geoscience. How to formulate foundation models tailored to geoscience, with implications for diverse downstream tasks, remains an underexplored area. Furthermore, humanity's quest for knowledge has increasingly extended beyond Earth into outer space.[72,73] The 21st century has seen significant advancements in space exploration.[74–76] For example, NASA's Artemis campaign aims to explore the Moon for scientific research and technological advancement in 2024,[77] alongside China's Chang'e program.[78,79] The BepiColombo mission of European Space Agency targets perplexing questions about Mercury, aiming to unravel the history of the entire Solar System.[80] With our knowledge of other planets still limited, advanced AI techniques play a crucial role in processing and analyzing the vast amounts of data collected from these missions. By deepening our comprehension of planetary processes, we cannot only enhance our understanding of these celestial bodies but also enrich Earth-based research by drawing insightful comparisons between fundamental geological mechanisms and planetary evolution.[81,82]

Advanced AI techniques, particularly emerging paradigms such as physics-informed ML and large models, showcase unprecedented potential for advancing geoscience. These innovative approaches open new avenues for addressing complex challenges not only in Earth science but also in the exploration of outer space. However, current research in these promising domains remains relatively limited. This article aims to offer a comprehensive overview of the latest advancements in AI applications within geoscience. In addition, it discusses the associated challenges and identifies untapped opportunities in this field, providing guidance for future works. While several reviews have previously explored the application of AI in geoscience, offering valuable insights into the evolving landscape,[59,61,70,83] the rapid evolution of AI, up-to-date reviews to capture current trends and illuminate future research directions. Geoscience, in particular, requires special considerations for AI methodology design, given the unique characteristics of geo-data. Therefore, rather than revisiting fundamental concepts and exemplified applications of commonly used ML models,[59,70,83] our work highlights the latest achievements and prospects of AI, especially in handling big geoscience data. We will demonstrate how AI can overcome the trade-off between efficiency and accuracy, as well as make breakthroughs in other aspects, such as providing new plausible hypotheses and research directions. Furthermore, we summarize new emerging geoscientific questions and paradigms in the context of modern AI and contemporary space exploration to shed light on potential future avenues for geoscience researchers.

The rest of the paper is organized as follows. The section "geoscientific research paradigms" summarizes major geoscientific research paradigms, with a special focus on AI-related ones in section "AI-driven geoscience paradigms" and some typical application cases in section "typical cases". The latest progress of geoscientific large models is demonstrated in section "large models in geoscience". Then, we present some challenges and plausible future lines for contemporary AI geoscientific method design in section "challenges and outlooks in AI for geoscience", followed by some findings in the "conclusion" section.

GEOSCIENTIFIC RESEARCH PARADIGMS

Diverse approaches and paradigms have been developed to deepen our understanding of the dynamic Earth system.[84] This section offers a

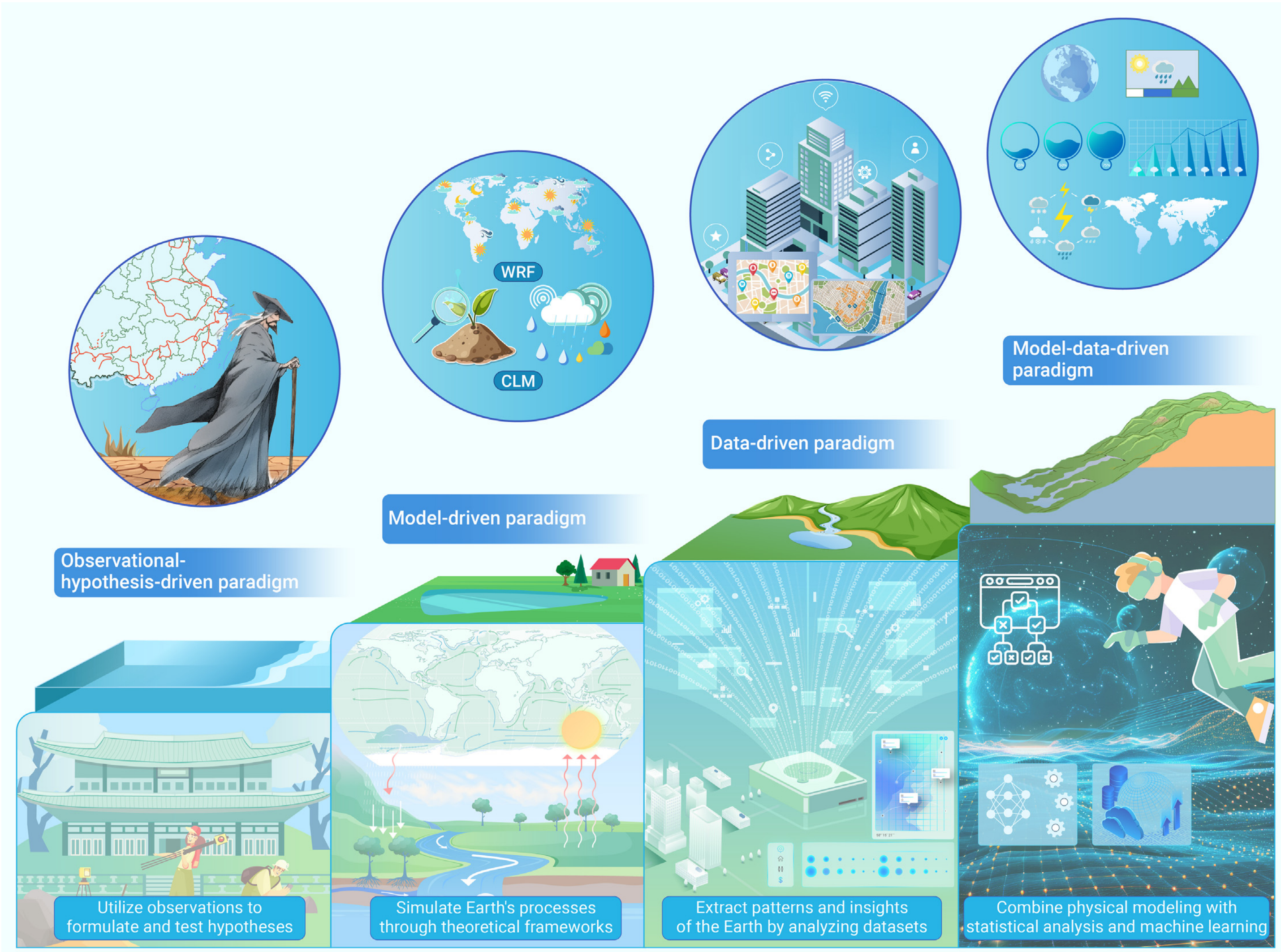

Figure 1. Illustration of four research paradigms in geoscience

comprehensive overview of the field, encompassing research paradigms ranging from traditional observational studies to advanced computational analyses. Four distinct yet interconnected methodologies have shaped contemporary geoscience: the observational-hypothesis-driven paradigm, the model-driven paradigm, the data-driven paradigm, and the model-data-driven paradigm,[10,18] as illustrated in Figure 1. Each of these methodologies brings unique strengths from foundational theories to advanced simulations and analyses, contributing to our comprehension of the Earth's complex system through collaborative synergies.

Observational-hypothesis-driven paradigm

The observational-hypothesis-driven paradigm is foundational to Earth system science, playing a crucial role in understanding the complex interactions within our planet's interconnected systems.[19,85] This approach involves systematic data collection and analysis to develop hypotheses about Earth's processes, dynamics, and derived consequences. Rooted in empirical evidence and scientific methods, this paradigm emphasizes objective observation and rigorous hypotheses testing.[86] A seminal example of this paradigm is James Lovelock's Gaia hypothesis,[87] which suggests the Earth's biosphere functions as a self-regulating system, a concept that fundamentally requires extensive Earth system observations to be substantiated. Observations validate and refine hypotheses, providing insights into processes that may not be directly observable. For instance, the study of ocean circulation has greatly benefited from observations of sea surface temperatures and currents, which have been crucial in understanding the dynamics of events such as *El Niño*.[88] In addition, using empirical observations and hypothesis testing in frameworks such as the Community Earth System Model has also been instrumental in assessing future climate scenarios and informing policy formulation.[89]

The advancement of technology has revolutionized our ability to collect data from various sources, including satellites, ground-based sensors, and remote sensing instruments.[90] Acquiring high-quality observational data has allowed scientists to refine their hypotheses and models, leading to more accurate predictions and a deeper understanding of Earth's behavior. In climate science, the Intergovernmental Panel on Climate Change Assessment Reports exemplify this paradigm in action. Leveraging extensive observational data, these reports critically evaluate the current state of climate system, hypothesize about future climate trends, and predict potential global and societal impacts. This demonstrates the profound impact of systematic observations on both scientific and policy-oriented discourse.[91]

In summary, the observational-hypothesis-driven paradigm is a fundamental method that combines empirical observations and hypothesis testing to unravel the interactions within Earth's interconnected systems. Firmly rooted in the scientific method, this paradigm remains indispensable for deciphering the intricate operations of the Earth system and guiding our responsible stewardship of the planet.

Future work within this paradigm should focus on advancing the integration and resolution of sensor networks across diverse ecosystems. By enhancing data collection methodologies, researchers can improve the accuracy of environmental models, leading to a more refined understanding of Earth system dynamics and their implications for global climate patterns. This approach will enable more precise predictions and foster a deeper scientific understanding of interconnected planetary systems.

Model-driven paradigm

There has been a long-standing focus on deciphering the interactions between natural processes and human activities on the Earth's surface.[92] This focus has driven the development of computational techniques and mathematical models, particularly process-based ones,[93] which simulate the physical, chemical, and biological processes of the Earth.[94] These models vary in complexity, ranging from simple representations of single processes to intricate integrations of multiple systems. The crux of process-based modeling lies in its challenges to transform our conceptual understanding of Earth processes into quantifiable and replicable frameworks.[95] By employing mathematical representations of natural phenomena, these models provide insights into the mechanisms driving Earth's systems.[96] Examples of such models include the Soil and Water Assessment Tool[97] and Storm Water Management Model[98] for hydrological studies, and the Finite Volume Community Ocean Model[99] for oceanic processes. These models have significantly enhanced our understanding and predictive capabilities regarding natural phenomena. In atmospheric science, models such as the Weather Research and Forecasting model,[100] Community Multiscale Air Quality model,[101] and Model of Emissions of Gases and Aerosols from Nature[102] are particularly pivotal. The predictive power of process-based models is substantial, allowing scientists to explore "what-if" scenarios that inform decision-making in environmental management and policy.[103] However, the efficacy of these models depends on their calibration and validation against empirical data.[104] This iterative process of refinement and validation ensures the models' accuracy and relevance, highlighting the continuous evolution of our understanding of Earth's systems through scientific inquiry and computational innovation.[105,106]

Future efforts in the model-driven paradigm should concentrate on refining model scalability and resolution, particularly by incorporating adaptive algorithms that improve the fidelity of simulations under varying climatic and environmental conditions in current and future scenarios. This will expand our capacity to predict subtle changes within Earth's systems with greater precision.

Data-driven paradigm

The data-driven approach has revolutionized our understanding of Earth systems and human-environment interactions.[107] Fueled by the vast availability of data and advancements in computing and sensing technologies, this paradigm allows researchers to gain deeper insights into the complex interplay between natural processes and human activities.[108] In geoscience, this paradigm shift is exemplified by utilizing satellite imagery and all kinds of big geo-data.[109,110] For instance, the analysis of observational data has enabled researchers to monitor changes in land cover, deforestation rates, and urban expansion, providing crucial information for sustainable land-use planning and climate change.[111–113] Data-driven methods have also transformed our understanding of urban environments.[114,115] The analysis of transportation data, such as traces from global positioning systems and traffic flow data, has enabled researchers to model urban mobility patterns and reduce traffic congestion.[116] In addition, social media data and geotagged content have provided insights into human behavior, sentiment, and urban cultural dynamics, shedding light on the social aspects of urban life.[117] The data-driven paradigm has also facilitated the study of the human-environment nexus in urban areas. By integrating data on air quality, land use, and human activity, researchers can better comprehend how urbanization affects air pollution, public health, and carbon neutrality.[118,119] This holistic approach has been instrumental in shaping policies aimed at improving urban air quality and reducing pollution-related health risks. Moreover, the integration of socioeconomic and environmental data has enhanced our understanding of urban resilience and vulnerability to natural disasters.[120,121] For instance, by analyzing demographic data and flood risk maps, researchers can identify vulnerable populations in flood-prone areas and devise targeted disaster preparedness strategies.[122]

In summary, the data-driven approach has propelled our understanding of Earth systems and urban dynamics to new heights. By harnessing vast datasets and sophisticated computational techniques, researchers can now explore the intricate connections between natural processes and human activities, facilitating more informed decision-making in areas such as land use, climate adaptation, transportation planning, and disaster resilience. This paradigm shift advances our scientific knowledge and offers practical solutions to the challenges facing our planet and urbanized societies.

Future work in the data-driven paradigm should emphasize the development of real-time data processing and analytics frameworks. By enabling instantaneous analysis and application of Earth system data, researchers can deliver more timely responses to environmental changes and disasters, thereby enhancing decision-making processes in critical situations.

Model-data-driven paradigm

The integration of process-based and data-driven models, commonly referred to as hybrid models, leverages the strengths of both paradigms and advances our comprehension of Earth system dynamics.[17] Hybrid modeling enhances simulation precision and computational efficiency.[123] Process-based models, underpinned by equations of 171 motion, are particularly effective in capturing the processes of atmospheric and oceanic dynamics. However, they often struggle with complex areas such as biological processes and carbon cycling, where numerical methods fall short and semi-empirical methods lack the necessary details and accuracy.[124] Hybrid models address this gap by employing ML to replace empirical sub-models, utilizing extensive observational data while maintaining process-based models for well-understood mechanisms.[125] In addition, certain components of Earth system models are computationally expensive, particularly when handling large datasets involving complex partial differential equations[126] or high-dimensional problems.[127] Despite the fact that ML emulators may incur high initial training costs, they offer a significant reduction in computation time once operational, outperforming traditional local process modules.[128] This increase in computational efficiency not only accelerates model processing but also enhances sensitivity and uncertainty analyses. The data-driven aspects of these hybrid models afford the flexibility needed to adapt to evolving conditions, as seen in climate and vegetation dynamic modeling.[17] Moreover, integrating physical principles into ML models enhances interpretability and extends their ability to extrapolate beyond observed datasets. For instance, domain-specific knowledge and models can be used to create synthetic data[129] or to select representative training samples,[130] which can train ML models that are both generalized and cost-effective. Unique neural network architectures that incorporate physical constraints, known as physics-informed neural networks, provide solutions to partial differential equations used in climate dynamics modeling.[131,132] In addition, embedding physical laws into the cost functions of neural networks, traditionally optimized by statistical measures such as cross-entropy or mean-square error, introduces a regularization effect and inherently discards physically implausible outputs.[133] The synergy between ML and physical modeling not only fortifies model credibility but also establishes a methodological evolution.

Future initiatives within the model-data-driven paradigm should concentrate on enhancing the scalability and integration of hybrid models across various scales and systems. This would include fine-tuning the interoperability between ML algorithms and process-based models to ensure seamless functionality in both regional and global-scale simulations. Such advancements could drastically improve the capability to simulate complex Earth system interactions and provide more accurate forecasts under changing climatic conditions.

This section has delved into the diverse paradigms and methodologies of geoscience, highlighting the multifaceted approaches to understanding our planet. The observational-hypothesis-driven paradigm forms the basis for empirical investigation, setting the stage for further inquiry, while the model-driven and data-driven approaches offer advanced simulation and in-depth analysis tools. In summary, the current landscape of research in geoscience has encountered limitations in effectively addressing complex global challenges.[59] There is a need for a transformative shift toward insights that integrate advanced AI techniques with geoscientific knowledge.[123] As geoscience continues to evolve, the interplay of these methodologies will be instrumental in driving forward our global efforts for environmental protection and sustainable development.[10]

AI-DRIVEN GEOSCIENCE PARADIGMS

Earth science research has undergone a transition from the observational-hypothesis-driven paradigm (see Figure 2) to a joint process-data-driven paradigm, which exhibits the characteristics of the "four Vs" of big data: volume, variety, velocity, and value.[134] Since the early 2010s, the performance of AI has improved dramatically[70] due to the availability of large-scale datasets, massive computer and storage hardware, and efficient distributed and parallel computing frameworks. The rise of AI has greatly accelerated the paradigm transition in

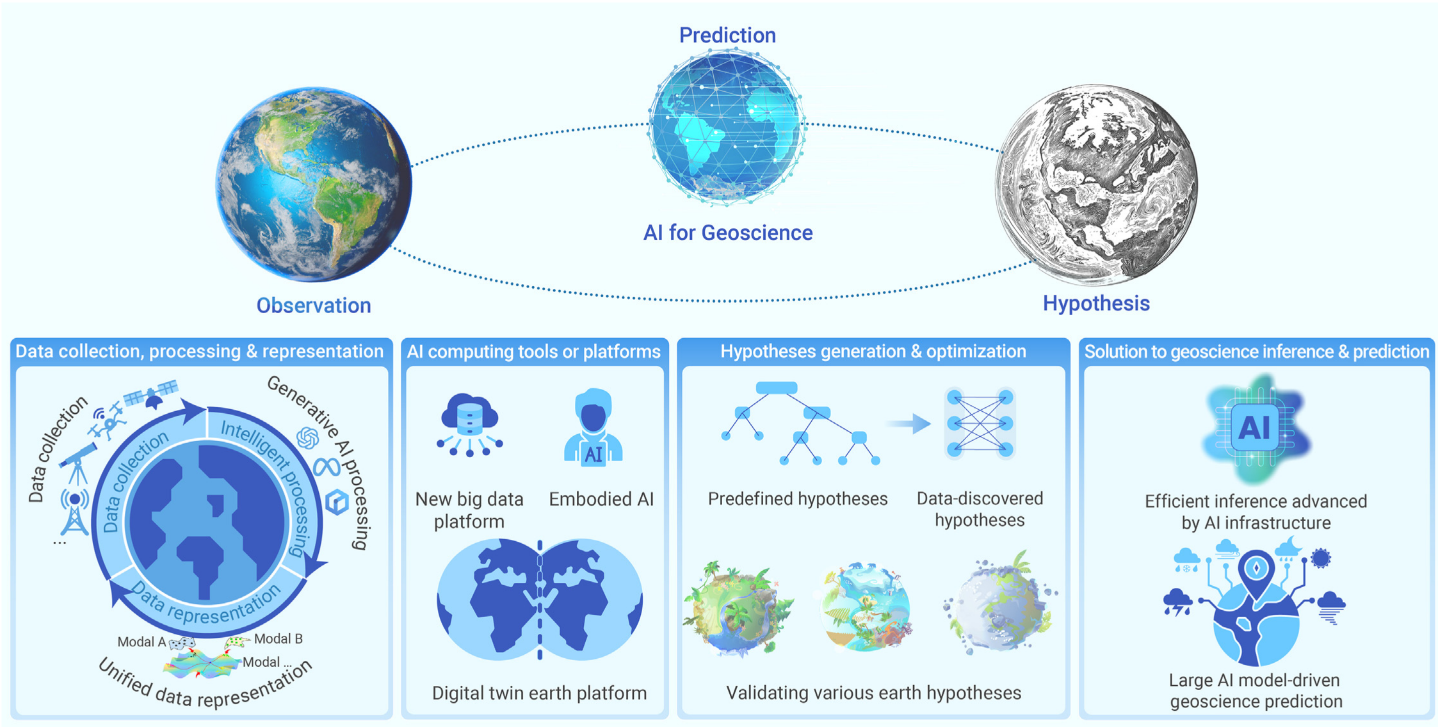

Figure 2. AI-assisted observations, hypotheses, and predictions of geoscience

geoscience research and driven various aspects of the application processes of big Earth data, from big Earth data collection and processing[135] to novel computational platforms,[136] hypothesis generation,[137] and geoscience prediction.[138] In this section, we discuss how AI can contribute to geoscience research in these aspects and the unprecedented opportunities it presents.

AI-assisted Earth observation data collection, processing, and representation

Data collection and analysis form the foundation of Earth science discoveries, aiming to capture, process, and represent complex Earth data to mine valuable information to understand complex Earth systems.[17,84] AI enhances and accelerates each stage of this process. AI accelerates and improves the efficiency of Earth observation data collection. The conventional satellite-to-ground data collection process typically includes multiple stages, requiring high time consumption and bandwidth.[139] Edge computing with AI on satellites allows real-time data processing and selective transmission to ground stations, significantly improving efficiency and reducing the need for manual corrections. Similar applications include real-time geographic information services for mobile terminals,[140] high-precision monitoring of ground stations,[141,142] and UAV-based agricultural remote sensing.[143] For example, Wang et al.[144] proposed a cloud-edge-end collaborative system for agricultural remote sensing, allowing AI to perform real-time data collection and processing on edge UAV devices. The processed data are then sent to the cloud, enhancing data transmission rates. In the future, integrated data collection and processing on edge sensors via AI[145] will be of great potential.

AI significantly contributes to data generation, completion, and enhancement. Earth observation data frequently encounter limitations in temporal, spatial, or spectral dimensions due to meteorological conditions, noise interference, and sensor issues, resulting in discontinuities across these dimensions.[109] Generative AI's ability to process multi-modal data across time, space, and multiple spectra is essential for generating, completing, and enhancing geoscientific data.[146] For instance, Kadow et al.[147] developed an AI model using inpainting technology to reconstruct meteorological data, restoring the missing spatial pattern of the El Niño event from July 1877. Moreover, large diffusion models such as DiffusionSat[148] and CRS-Diff[149] are capable of performing integrated tasks and addressing the issue of limited remote sensing samples in specific spatiotemporal scenarios. SpectralGPT[150] captures spectral sequence patterns through multi-objective reconstruction, providing a foundation model with over 600 million parameters for various downstream tasks.

AI enhances the flexibility and effectiveness of data representations by introducing geometry and structure to model the complex interrelations within the data.[59,151] For example, graph networks[152,153] model directly underlying structures, facilitating the discovery of broader spatial correlation patterns in Earth science data. Self-supervised learning[154] allows capturing general features without relying on explicit labels. The Transformer architecture,[155] known for its powerful feature extraction and long-distance spatial dependency modeling capabilities, unifies data representations across various scenarios and modalities in Earth science. Recently, large AI models have revolutionized representation learning by facilitating deep interconnections between Earth science data to unearth new scientific discoveries. Examples include the single-modal large language model K2,[156] the meteorological time series graph network model GraphCast,[157] and the large multi-modal model SkySense, which integrates images, text, geographic coordinates, and site observations.[158] Exemplified by digital twin Earth, unified and universal large Earth science models have become a future trend.[150,159] Their embedding representations should not only consider the capabilities of multi-scale spatiotemporal data processing, multimodal data representation, and alignment with human understanding, but also provide a universal interface for decoders tailored to various downstream tasks, achieving comprehensive generalization across the domain.

AI promoting new computing tools or platforms for geoscience research

Numerous processes on and within the Earth are continuously monitored by various sensors globally, generating vast amounts of Earth data, with storage volumes exceeding 10 exabytes.[19,160] These sensors capture various states, fluxes, and intensities, capturing time/space-integrated data from satellite remote sensing, *in situ* observations, and atmospheric monitoring devices.[19] Traditionally, geoscientific systems have required the integration of decentralized decoders tailored to specific tasks to compute and simulate the diverse and spatiotemporally varied streams of observational data. This approach often complicates data sharing and model connectivity. However, the emergence of new AI tools and large models is poised to revolutionize the computation and simulation paradigms of geoscience big data platforms.

First, large AI models are driving the innovative construction of big data platforms that offer robust multi-task processing capabilities and efficient data integration mechanisms. For example, AI-Earth[161] introduced AI-Seg, a universal foundation model for object segmentation, capable of rapidly segmenting multi-source remote sensing images and extracting spatiotemporal change information. The Open Geospatial Engine[162] incorporates LuojiaNet, a DL architecture tailored to geoscientific features, linking 55 downstream foundation models with 300 million parameters. This system also includes an embedded spatiotemporal knowledge graph to associate multimodal spatiotemporal data.

Second, AI agents, in collaboration with high-quality feedback from geoscience experts, can assist in solving complex geoscientific processes or problems. The "human-in-the-loop" process, which involves deep models and human experts, has proven effective in geoscientific data annotation with improved interpretation accuracy.[163] For instance, Li et al.[164] have integrated large conversational models into robots, allowing humans to issue commands to robots via language for complex action-planning tasks. This advancement in AI's understanding of spatial intelligence is catalyzing robotic learning, approaching the goal of embodied intelligence.[165] Moreover, the deep integration of large AI models with drones, autonomous vehicles, and mobile monitoring devices on the surface or underground, coupled with satellite data, is facilitating more efficient and automated complex actions in geoscience, including spatiotemporal data collection, processing, and transmission.

Finally, the integration and assimilation of Earth's big data through the digital twin Earth has ushered in a new era of experimentation and simulation in Earth science.[166] By integrating remote sensing data, *in situ* observations, experimental analyses, societal perceptions, simulations, and reanalysis, AI-based digital twin systems or platforms are capable of accurately simulating various complex Earth processes, spanning atmospheric, hydrological, urban, geological, and other domains.[167] Specifically, Earth digital twins, which integrate big Earth data and physics-based models within interactive computational frameworks, enable the monitoring and prediction of environmental changes and societal disruptions,[168] thereby driving a deeper understanding of Earth system processes and scientific cognition.

AI facilitating the generation and optimization of geoscientific hypotheses

Hypotheses are crucial research tools in Earth sciences, aiding scientists in comprehending the Earth system and its evolution through artificial observations and scientific conjectures.[169] For instance, Kepler[170] formulated the laws governing planetary motion based on extensive observations of stars and planets. Geoscientific hypotheses appear in various forms, including mathematical expressions, molecular formulas in geochemistry, and genetic variation laws in biology. Traditional methods for generating and validating hypotheses have predominantly relied on theoretical assumptions and logical deduction, as well as computational modeling and simulation,[18] with limited ability to solve complex and nonlinear problems. In contrast, recent AI has learned patterns and rules in massive data through "guessing-and-verifying," with intelligence gradually emerging.[171] This evolution has led to significant breakthroughs in scientific endeavors, such as predicting protein structures,[172] formally proving mathematical conjectures and theorems,[173] and simulating molecular dynamics in physics.[174] This "guess-and-verify" type of AI has greatly contributed to the paradigm shift of geoscientific hypotheses generation and validation.

AI is transforming the generation of geoscientific hypotheses from predefined methods to data-driven discovery. Traditionally, hydrologists modeled rainfall-runoff processes using physical conceptual models based on potential influencing factors,[175,176] which tend to be non-unique, subjective, and limited.[177] In contrast, AI treats multimodal data as inputs, enabling scientists to explore larger sets of hypotheses for more effective generation.[22,178] Furthermore, screening a high-quality hypothesis from the candidates is usually framed as an optimization problem.[179] AI prioritizes directions with higher values by maximizing reward signals for the candidate set, instead of using manually designed rules in the traditional approach.[180,181] For example, a multi-objective optimization framework was constructed to consider the impacts of hydropower capacity on five environmental factors (sedimentation, river connectivity, flow regulation, biodiversity, and greenhouse gases) in the Amazon basin.[182] AI also enables selective screening of candidate information with desirable attributes from high-throughput experimental data, reducing the interference of redundant observations.[183] Another example is the optimization of discrete geo-hypotheses, where AI methods, such as variational autoencoders, map discrete symbolic representations into a differentiable latent space.[184] Process-based differentiable modeling[63] combines physical mechanisms and ML techniques, facilitating hypothesis testing and uncovering previously unrecognized correlations in Earth science.

AI holds the potential to significantly contribute to the verification of geoscientific hypotheses. Various hypotheses in geoscience, such as Wegener's continental drift theory, Darwin's biological evolution hypothesis, and the historical climate change conjecture, present major scientific challenges. Correspondingly, researchers have leveraged AI's ability to model nonlinear complex systems to verify these hypotheses. For example, Stupp et al.[185] used coevolutionary ML to predict functionally relevant interactions between human genes, advancing the understanding of human coevolutionary processes. Kalra et al.[186] utilized artificial neural networks to model the complex association between global temperature and greenhouse gas concentrations. In addition, large AI models such as GraphCast[157] and PanGu[187] have revolutionized traditional weather forecasting methods and contributed to exploring the evolution of Earth's climate over deep time. AI also challenges the findings of traditional physical models.[17] For example, ML estimates of global carbon flux data have indicated that traditional climate models may have overestimated the response of vegetation, such as tropical rainforests and grasslands, to climate changes.[188] Data-driven carbon cycle estimates have also revealed potential mechanisms behind the enhanced seasonal cycle of atmospheric carbon dioxide concentration in high-latitude regions.[189]

AI-driven solutions to geoscience inference and prediction

Geoscience prediction tools have undergone substantial evolution, improving our ability to comprehend complex Earth systems.[190] Initially, Galileo, Kepler, and others studied planets through experimental methods of observation and induction.[191] Alfred Lothar Wegener studied the Earth's plates through hypothesis and deduction.[192] Subsequently, the simulation and modeling of complex phenomena, such as meteorological, hydrological, oceanic, and other physical processes, through physical computational models became the third paradigm of Earth science research. With the arrival of the big Earth data era and the continuous improvement of AI, scientists have made significant strides in spatiotemporal analysis.[151,193] The data-intensive research paradigm has become a mainstream.[11] Nowadays, large AI models have revolutionized the paradigms for geoscience inference and prediction,[17] exhibiting strong abilities to mine hidden relationships within vast data and enhanced model inference and prediction accuracy.[194,195]

AI allows for more comprehensive and efficient geoscience inference and prediction. To enhance geoscience inference, AI implements trustworthy attention-based models, enabling the extraction of spatial relationships across data from a global perspective.[196] Physical embedded neural networks leverage their powerful numerical approximation capabilities to reduce the computational complexity of high-order differential equations.[197] Spatiotemporal graph neural networks, which utilize the graph structures to accurately represent spatial relationships and factor correlations, enhance the reliability of reasoning.[195,198] In terms of prediction, AI incorporates a broad range of historical information and efficient modeling strategies, such as pre-training[199] and generative decoders,[200] offering enhanced technical support for decision-making processes. With the deep integration of AI infrastructures (such as high-performance computing chips, storage media, rapid and lightweight large models) and edge sensors, real-time monitoring of the Earth's environment contributes to enhancing the predictive capabilities for rapid disturbances such as geological disasters, climate anomalies, and emergencies.[91,154] Overall, AI facilitates more precise and reliable inference and prediction, reducing the computational complexity of high-order differential equations.[201,202]

Society has witnessed many successes in this respect, although many challenges still exist. Weather prediction, a successful example in geoscience, has dramatically improved through integrating advanced AI models, increased computational power, and established observational systems with large amounts of data.[122] Represented by PanGu[187] and GraphCast,[157] large AI models can accurately predict weather evolution on time scales ranging from several days to a month. However, challenges remain in seasonal weather forecasts, extreme event predictions (such as floods and wildfires), and long-term climate forecasts.[120,203] In the biosphere, Klemmer et al.[204] trained a universal AI geolocation encoder to assist in monitoring biological population migration and number estimation. However, dominated by biologically mediated processes

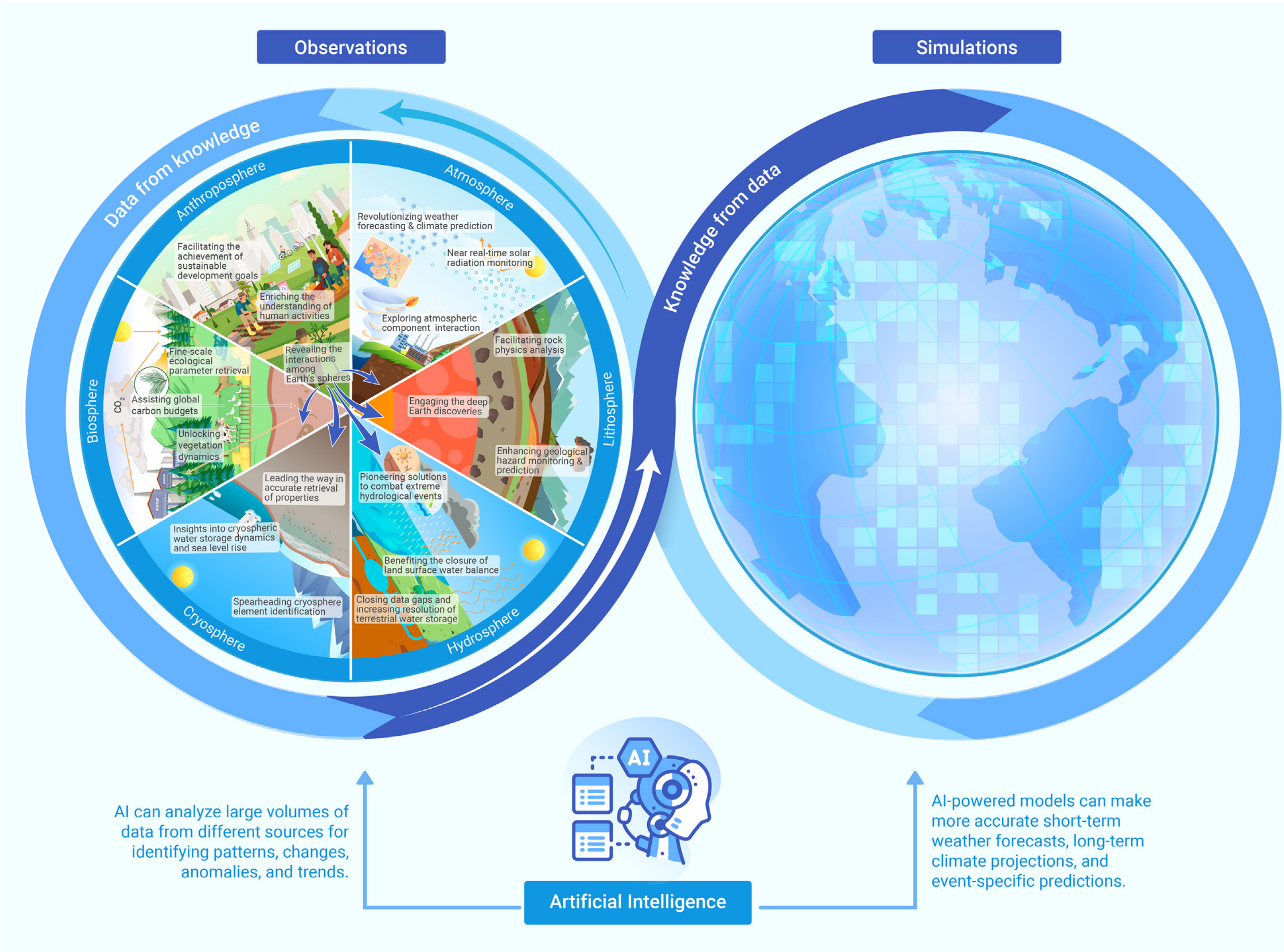

Figure 3. Observation and simulation are the two main tools for understanding the Earth system AI helps in the observation of the Earth system, assisting in the discovery of knowledge from data. Besides, AI also supports Earth system simulation, generating data from models and knowledge.

such as reproduction and migration, and influenced by seemingly random but intense disturbances such as earthquakes, landslides, and volcanic eruptions, predicting the dynamic changes and deep-time evolution of the biosphere remains difficult.[205–207] It is also essential to establish a comprehensive, integrated monitoring network in outer space, sky, surface land, and subsurface to provide more reliable data support.

TYPICAL CASES

AI, as a modern scientific research infrastructure that comprises rapidly evolving technologies, brings novel means to comprehend the Earth's systems, including the atmosphere, lithosphere, hydrosphere, cryosphere, biosphere, and anthroposphere, as well as their interactions. By leveraging rapidly advancing technologies, AI accelerates and deepens our understanding of the Earth at a variety of spatial and temporal scales, advancing the achievement of sustainable development goals (see Figure 3). The uniqueness of geoscience, showcasing a considerable amount of subdisciplines, a vast quantity of geographic knowledge, an extensive collection of observational data and spatial dependence, spatial heterogeneity, and nonlinearities among geographical elements, has led to novel advancements in AI technology.

Atmosphere

Clouds, aerosols, and gases are three of the most important components in the atmosphere. They affect the solar radiation received by the Earth system and exert distinct radiative forcing on the energy budget, which in turn has a substantial influence on the weather and climate on a regional or global scale.[18,208,209] AI models the complex and nonlinear atmosphere system, predicting common surface and atmospheric variables, as well as enhancing our ability to retrieve atmospheric parameters with remarkable enhancement in the accuracy and granularity of atmospheric studies.[210,211]

Atmospheric component detection and interactions. AI has revolutionized cloud identification, cloud type recognition, and even cloud dynamics prediction from satellites,[212] It has notably improved the accuracy of retrieving cloud microphysical and cloud top parameters[213–215] and has provided cloud bottom information that traditional physical-based algorithms often fail to estimate.[216] These advancements enable the precise understanding of cloud formation in weather forecasting,[217] holding the promise of more accurate and efficient weather predictions.

In aerosol remote sensing, AI mainly contributes to improving the detection of aerosols,[218] building models to retrieve aerosol optical properties[219–221] and applying the aerosol products to wildfire detection, particulate matter ($PM_{2.5}$) monitoring, and other aerosol-related problems.[219] It is noteworthy that AI is becoming an irreplaceable tool to develop high spatiotemporal datasets of the aerosols originating from various emission sources that improve our understanding of the climatic, environmental, and health effects from the intricate composition of aerosols.[222–224]

AI models, on the one hand, retrieve water vapor with high accuracy[216] and produce precipitation datasets with a high spatial and temporal resolution.[225] On the other hand, AI techniques have been integrated with ground- and satellite-based observations to quantify and forecast air quality on a regional or global scale.[216] Many factors (meteorology, geography, emissions, vegetation, etc.)

have been incorporated into the AI models to explore the complex non-linear relationships between satellite-based observations and the surface concentrations of various gaseous air pollutants, which provides insight into developing more efficient strategies to reduce the adverse health and societal effects of air pollution exposure.[226–229]

By advancing cloud analysis, improving aerosol monitoring, and exploring the relationships between complex gases, AI significantly enhances researchers' understanding of the dynamics of atmospheric components and captures their intricate interactions.

Solar radiation monitoring. The traditional radiative transfer (RT) model is a classic and widely used way to retrieve solar radiation. However, the forward RT simulation is a time-consuming process, which makes it inapplicable for direct use with satellite observations, particularly with geostationary satellite observations (the monitoring frequency in the order of minutes). Through the development of AI-based RT models in recent years, the computational efficiency of atmospheric RT has been greatly improved (by several orders of magnitude),[230] which enables near real-time monitoring of solar radiation from satellites with high accuracy.[231]

Weather forecasting and climate prediction. Mainstream AI-based global weather/climate forecast models predominantly concentrate on short- and medium-term predictions,[151] such as Google DeepMind's GraphCast,[157] Huawei Cloud's Pangu-Weather,[187] Tsinghua University and China Meteorological Administration's NowcastNet,[232] Alibaba's SwinVRNN,[233] Fudan University's Fuxi,[234] Shanghai's AI Laboratory's Fengwu,[235] Microsoft and the University of Washington's Deep Learning Weather Prediction,[236] with exceptional capabilities in processing large datasets, performing real-time analysis, and predicting extreme weather events.[237] AI-based global weather/climate forecasting models have high forecast timeliness and computational efficiency. Taking Pangu-Weather as an example, it predicts 7 days' weather in only 10 s, 0.6 days earlier than the world's leading weather forecasting system, the European Center for Medium-Range Weather Forecasts (ECMWF).[187] It is of great significance for extreme weather forecasting. Based on the weather forecast assessment of China's National Ground Meteorological Stations in the first quarter of 2024, AI-based models such as Fuxi, GraphCast, and FourCastNet had higher accuracy in temperature and wind speed than traditional numerical predictions.

Atmospheric predictability revolution: From challenges to solutions. The atmosphere is an intricate and dynamic system, and myriad challenges originate from the subtle interaction among aerosols, clouds, gases, and radiations.[210] Predicting weather patterns and understanding climate change accurately are paramount in atmospheric science. However, achieving these goals poses significant challenges, including the need for faster and more precise weather forecasting and climate projections. Nowadays, AI models have emerged as a powerful tool for tackling these challenges and advancing solutions across a wide array of applications in atmospheric sciences.[237] It significantly promotes the development of related monitoring and prediction platforms, which produce massive data and information with high spatiotemporal resolution and improved accuracy.[10,234] In the future, as AI continues to evolve and incorporate more spatial big data into its training, it will enhance the reliability and accuracy of weather and climate forecasts further. Consequently, it may even lead to the eventual replacement of traditional physics-based models with AI-driven approaches. In addition, AI will play an imperative role in constructing automated monitoring and warning systems for the atmosphere environment, enabling timely issuance of alerts and recommendations. In essence, AI's application in atmospheric science transcends traditional methods, providing innovative solutions to long-standing challenges. Its integration into timely and accurate monitoring and prediction systems not only advances our understanding of atmospheric processes but also empowers us to make well-informed decisions.

Lithosphere

Solid Earth science, aimed at comprehending the structure, materials, and dynamics of the Earth's interior, geological processes, and the evolutionary history of the Earth,[238] receives unprecedented opportunities from AI,[10] with dramatic developments in geological hazard monitoring and prediction, rock feature analysis, geological exploration, geological model construction, and analysis of soil characteristics.[10,239]

Geological exploration and hazard prediction. AI approaches have made significant strides in their application to geological exploration, such as petroleum and natural gas exploration,[240] geophysical imaging,[241] as well as the processing of seismic,[242] magnetotelluric,[243] and gravity data,[244] enabling sophisticated analysis and interpretation. Techniques such as denoising, phase-picking, and weak signal enhancement can reduce human errors in the exploration process, enhance the quality of exploration data, and accelerate exploration time.[245] By harnessing the power of AI, geoscientists can unlock new frontiers in exploration efficiency, accuracy, and cost-effectiveness, ultimately shaping the future of resource exploration and sustainability.

AI technology provides powerful tools in facilitating earthquake monitoring and prediction, including detection and phase identification,[246] early warning,[247] motion prediction,[248] as well as forecasting magnitudes, scales, and timing,[249] also in assessment of landslide susceptibility,[250] supporting the mitigation of risks. AI has also demonstrated great potential in the volcanic prediction process.[251] Although AI encountered grand challenges in operational earthquake prediction, forecasting of fault zone stress, and the occurrence of chained natural hazards attributed to their highly coupled and strongly non-linear dynamics,[252] it has exhibited tremendous progress in recent years.[253]

Rock physics analysis. AI methods can be utilized for the analysis and classification of rock samples,[254] automatically identifying rock types, compositions, and physical characteristics, thus expediting the analysis of rock samples and providing more detailed information about rock features, so as to aid geologists in better understanding geological history and rock evolution.[255] Recent evidence demonstrates that AI has successfully solved various problems in rock mechanics, outperforming conventional empirical or statistical methods.[254] By using AI approaches, deeper insights can be gained for more accurate geological interpretations and predictive models.

Geological modeling. As the emerging paradigm of science and technology research, AI is modulating the world in a variety of science realms, including geology. AI is transforming the measures geologists analyze data and understand the mechanisms of deep Earth. During the construction of geological models, AI is capable of integrating vast amounts of geological data from various disciplines and fields, ranging from geophysics and geochemistry to hydrology and tectonics. This multidisciplinary approach generates predictive models that assist scientists in better comprehending subsurface structures, stratigraphic forms, and groundwater flow. AI-driven predictive modeling helps geologists to efficiently and accurately identify patterns and trends that were difficult to detect early on.[256–258] AI's integration into deep Earth modeling enables geologists to identify previously unrecognized geological features and phenomena, thereby advancing our understanding of the deep Earth.

Soil characteristics monitoring. AI boosts new developments in soil monitoring, offering a holistic and data-driven approach to soil monitoring and management.[259] AI-driven sensors and monitoring systems enable continuous and high-resolution monitoring of soil conditions, providing valuable insights into soil health and dynamics, including essential soil parameters such as moisture,[259] temperature,[260] and texture.[261] By analyzing multispectral and hyperspectral imagery, AI models can accurately map soil types, nutrient levels, and organic matter content across large spatial scales. This new real-time ability assists farmers in making informed decisions, thereby improving farmland utilization efficiency and agricultural production quality.[262]

Deep-time and deep-Earth discoveries: Scale and accuracy. To date, large AI models constitute the most cutting-edge and wisdom-intensive research regime; the integration of AI has indeed ushered in a new era of exploration and understanding. However, as we delve deeper into the complexities of lithospheric processes, it becomes apparent that simply scaling up large AI models without due consideration for their ability to accurately capture and resolve scientific intricacies may lead to deviations from fundamental physical laws and characteristics. While large AI models boast impressive computational power, their efficacy in accurately describing lithospheric phenomena may be limited by the uncertainties inherent to input labels and data. Therefore, a shift toward the development of numerous and accurate small-scale domain-oriented models tailored to specific scientific problems or application fields is warranted. In addition, the integration of high-quality observation, monitoring, and experimental data with completeness is crucial for training and validating AI models in lithospheric studies.[263] Synthetic data derived from massive-scale numerical simulations can further enhance the robustness and generalizability of AI models. Essentially, it may be a reliable and feasible measure to promote the revolutionary engagement of AI in deep-time and deep-Earth discoveries.

Hydrosphere

The hydrosphere is the sum of all water, including atmospheric, land surface, oceanic, and underground water reserved on Earth.[264] AI addresses a wide range of applications in the hydrosphere that allow (but are not limited to) better modeling and estimation of precipitation, soil moisture, evapotranspiration, streamflow, water storage, ocean currents, and ocean salinity, by simulating the complex input-output relationships inherent to nonlinear hydrological processes, thus improving the accuracy of hydrological model simulations and remote sensing retrievals.[265–267]

Land surface water balance. AI benefits the closing of the land surface water balance by accounting for the individual surface water flux components (precipitation, evapotranspiration, streamflow) and expanding the mapping capabilities of key state variables (such as soil moisture). AI can improve the estimation and forecasting accuracy of precipitation and help better understand the causes of extreme rainfall.[268] For example, generative adversarial networks have been used for precipitation nowcasting and proved to be of high reliability.[269] The multi-layer perceptron model, integrating geostationary satellite infrared data and passive microwave-based retrievals, yields precise precipitation estimates.[270] Probabilistic weather models such as deep neural networks (i.e., MetNet-2) forecast precipitation with exceptional resolution, up to 12 h ahead.[271] Moreover, AI empowers the generation of precipitation data with unparalleled precision, spatiotemporal resolution, and spatial coverage, enhancing our understanding of precipitation dynamics.[272] In addition, AI methods analyze large-scale circulation patterns associated with US Midwest extreme precipitation to better understand the physical causes of changing extremes.[273] Despite the many successful cases of AI application, acquisition of high-quality and continuous atmospheric data is still challenging due to sensor limitations, and the implementation of hybrid models appears as an effective solution.

AI-based approaches have been extensively employed to estimate evapotranspiration, one of the most important components of the hydrological cycle. That is crucial for estimating irrigation water requirements, hydrological processes, and assessing agricultural systems at both regional[274] and global scales.[275] Site-scale evapotranspiration observations can be upscaled to the regional scale using AI-based methods,[276] thus overcoming the limited spatial and temporal coverage of *in situ* observations. The ability of AI to forecast evapotranspiration is also highlighted in a recent study.[277] These forecasts play a crucial role in agricultural planning and drought monitoring, contributing to improved resilience and sustainability in water management practices. A novel research direction is to estimate evapotranspiration at high resolution through the construction of hybrid models,[278] which combine the physical consistency and interpretability of physical models with the data-driven formulations of AI-based models, thereby revealing processes that are insufficiently understood. This interdisciplinary approach holds the potential to uncover the underlying mechanisms and diversity of evapotranspiration, thereby enabling more robust and insightful assessments of water cycle dynamics.

Streamflow, as a key aspect of sustainable water resource planning and management, can be estimated in real time[279] or forecasted at lead times of 1–7 days[280] AI-based approaches, successfully used in streamflow regionalization,[281,282] can help to reduce modeling errors in process-based hydrologic models to improve the accuracy of simulations, since process-based and AI approaches can complement each other with respect to their inherent strengths and limitations.[283] Deep neural networks enable the accurate identification of spatial distribution and morphological features of water bodies,[284] understanding river evolution, and forecasting river dynamics,[285] performing water quality analyses on the catchment scale.[286] Another significant contribution of AI is the creation of global water quality databases due to its powerful learning and data fusion capabilities, such as the Global Streamflow Indices and Metadata Archive,[287] global river discharge reanalysis,[288] Global River Chemistry Database,[289] and Global River Water Quality Archive.[290] The integration of AI into streamflow estimation, forecasting, and water quality analysis offers transformative opportunities for strengthening our understanding of hydrological processes toward more sustainable and resilient water systems.

Soil moisture acts as a fundamental boundary condition in terrestrial hydrology.[178,291] The integration of AI-based models into soil moisture mapping significantly advances our ability to accurately retrieve, downscale, and predict soil moisture dynamics across different spatial and temporal scales. By using AI-based models, soil moisture retrievals are obtained from the passive-only[292] and synergistic active-passive microwave observations[293,294] with improved accuracy and temporal resolution, which is challenging for traditional algorithms to separate and interpret the desired information accurately. AI techniques downscale soil moisture from coarse spatial resolution to fine resolution,[295–299] also establish long-term global daily surface soil moisture datasets from multi-frequency radiometers (AMSR-E/2 and FY-3 series) by transferring the Soil Moisture Active Passive L-band observations, offering extended records vital for climate monitoring and hydrological research.[300,301] Moreover, with the help of AI algorithms, soil moisture can be predicted at deeper depth (e.g., root zone) from surface data[302,303] and in a seamless and efficient manner through AI-based data assimilation techniques.[304,305]

AI provides a potential solution for avoiding closure errors by enhancing the estimation and prediction of individual water fluxes and state variables. It addresses challenges related to integrating diverse data sources to produce cohesive models, achieving fine-scale spatial and temporal resolution, and understanding the nonlinear nature of hydrological processes.

Terrestrial water storage. AI plays an important role in improving the spatial and temporal continuity and resolution of terrestrial water storage. AI approaches have been instrumental in reconstructing continuous total water storage, by filling the data gap between the Gravity Recovery and the Climate Experiment (GRACE) satellite mission and its successor, GRACE-FO.[306] Similarly, AI-based models, such as the GTWS-MLrec, have been developed to reconstruct terrestrial water storage estimates spanning several decades from 1949 to 2022, using a set of ML models with a large number of predictors.[307] AI was used to capture complex spatiotemporal patterns in water storage dynamics, facilitating comprehensive analyses of hydrological trends and variability over extended periods. In addition, AI-based approaches have been deployed to map soil water storage in Ghana at high spatial and temporal resolutions, facilitating the identification of areas with stable water availability for improved crop production and guiding drought adaptation strategies.[308] Moreover, GRACE-derived terrestrial water storage anomalies are downscaled to 10 km spatial resolution by using a convolutional long short-term memory neural network[309] in Iran and convolutional neural network-based approaches in Canada.[310] In essence, AI provides a new capability to overcome data gaps, improve spatial resolution, and enhance the continuity of water storage observations, ultimately contributing to more effective water resource management.

Ocean currents and salinity. Ocean currents and salinity are crucial for understanding global climate systems, marine ecosystems, and coastal environments. Ocean currents reflect the movement of ocean water and drive the distribution of heat, nutrients, and salinity, influencing weather patterns, climate regulation, and marine biodiversity. AI methods significantly improve the estimation and forecasting of ocean currents by enhancing computational efficiency and accuracy.[311] Traditional methods often struggle with the complexity and volume of oceanographic data. AI models, such as those integrating sea surface height, temperature, and wind stress simulated from the ocean general circulation model, can accurately predict the ocean currents over most of the global ocean,[312] and successfully forecast the velocity. For structures of the loop current system,[313] AI techniques such as LSTM recurrent neural networks and the Transformer also enable real-time *in situ* prediction of ocean currents at any location, and overcome the problem of excessive computational complexity in traditional regional physics-based prediction models.[314]

AI-based approaches, such as deep neural networks, generative adversarial networks, random forests, support vector regression, and multi-layer perceptrons promote the convenient and fast estimation of ocean salinity, from the Aquarius,[315] SMAP,[316] and the Geostationary Ocean Color Imager-II satellites.[317,318] With the aid of AI-based methods, the ocean general circulation model (e.g., Hybrid Coordinate Ocean Model) is also able to achieve more reliable estimates of sea surface salinity.[319] AI has demonstrated strong capabilities to reconstruct the high-precision and high-resolution three-dimensional (3D) ocean subsurface salinity on a daily scale in 12 depth levels (from 2 to 200 m) only relying on the ocean 3D temperature data.[315] This is because AI, particularly the DL models, have flexible structures and can extract potential complex mappings of data by stacking only multiple nonlinear layers.

Extreme hydrological events: Pioneering solutions. The changing dynamics of global climate present two concerning trends in the hydrosphere: alterations in water circulation patterns and the increasing frequency and intensity of extreme hydrological events.[111,320] In response to these challenges, it becomes imperative to strengthen monitoring efforts, enhance forecasting capabilities, and improve decision-making efficiency. AI provides important tools for monitoring, understanding, and forecasting extreme hydrological events such as drought, rainstorm, and flood.[122] AI can integrate a large amount of data from various sources (e.g., satellites, meteorological stations, and other sensors) to provide more comprehensive and accurate monitoring results of extreme hydrological events.[321]

For example, for extreme events, Earth observation data and ML can significantly mitigate the scarce hydrological data. Satellite-based technologies, which encompass a wide array of sensors operating across different regions of the electromagnetic spectrum—such as visible, thermal, and microwave domains—offer considerable potential. Advanced sensors, including synthetic aperture radar (SAR), satellite-based precipitation measurements, and gravity measurements, are emerging as transformative tools for the forecasting and monitoring of extreme events.[322] Concurrently, the robustness and transferability of ML techniques are proving instrumental in predicting floods in ungauged river basins.[122]

Moreover, AI can analyze and learn from historical data and meteorological forcings (such as precipitation and temperature), and identify the interactions between different environmental factors, and thus help understand the causes and patterns of extreme hydrological events.[323] Furthermore, by using AI, short-term forecast of hydrological events can be made based on real-time hydrological data, providing timely support for emergency response. Short-term flood forecasting, which spans from a few hours to several weeks, predominantly utilizes meteorological forecasts to enhance model prediction performance and ensure physical consistency. For example, Xu et al.[324] have summarized numerous hydrological forecast models in this context. The prevailing methods for short-term flood forecasting integrate meteorological inputs (such as precipitation and temperature) with optional historical data to predict runoff or flooding events.

Meanwhile, combining meteorological and hydrological models, AI can forecast the long-term trend of extreme hydrological events, helping decision-makers to make long-term plans.[325] Long-term forecasting of extreme events, which includes sub-seasonal, annual, and decadal outlooks, remains a significant challenge due to inherent data and model uncertainties. Currently, hybrid learning approaches[324] that combine physical modeling with ML are being employed to reduce model uncertainties and mitigate the reliance of data-driven models on extensive data inputs. In addition, uncertainties in data (such as precipitation) can be addressed by integrating low-latency satellite observation data with reliable climate prediction models. In summary, AI has brought new opportunities for hydrological cycle research to better understand and cope with extreme hydrological events. With the rapid development of computer technology and the emergence of new interpretable AI methods, the role of AI in the hydrosphere (particularly in extreme hydrological events) will become more prominent in the future.

Cryosphere

The cryosphere refers to frozen components of the Earth system,[326] overlapping with the atmosphere, the hydrosphere, and the lithosphere over vast areas, exhibiting a sensitive response and holding a significant impact on climate change.[327,328] Numerous scholars focused on developing AI methods for addressing the challenging geoscientific questions in cryosphere research, such as the AI for Cold Regions, bringing new perspectives and innovative solutions in element classification and automatic mapping, feature spotting, physical properties retrieval, and interpretation of the cryosphere changes.[329]

Cryosphere element identification. AI overcomes ambiguity in the cryosphere element identification caused by feature similarity, superseding manual interpretation, and limited empirical approaches. One notable application of AI is that it enhances our comprehension of the spatiotemporal distribution of the cryosphere by better classifying its elements, such as distinguishing ice cover types,[330] especially debris-covered glaciers,[331] which were difficult for band ratios/indices. AI can overcome inherent complexities to generate high-resolution maps of permafrost, a critical component of the cryosphere.[332] Kuter et al.[333] applied artificial neural networks to estimate areal snow cover extent with high accuracy, thus able to provide timely and reliable information on snow cover dynamics. Convolutional neural networks have proven effective in classifying sea ice types with higher accuracy and less sensitivity to noise in SAR images.[334] Coincidentally, AI achieves automatic and reliable iceberg detection in different environmental conditions and improves understanding of iceberg dynamics in polar regions.[335] In essence, AI plays a significant role in spearheading our understanding of the cryosphere by overcoming traditional limitations in element identification.

Feature spotting. In addition to classifying the cryosphere elements, AI aids in the identification of specific features of these elements that were previously challenging to detect. Specifically, AI has advanced the identification of wet and dry snow, especially in vegetated and mountainous areas where traditional methods struggle to differentiate between snow types.[336] AI enabled robust and automated detection of snow avalanches for enhancing safety measures in mountainous regions.[337] In glaciological research, AI has been utilized to map glacier calving margins[338] as well as glacier termini[339] toward comprehensive assessments of glacier mass loss. Qayyum et al. developed a DL-based glacial lake extraction method with noteworthy benefits in monitoring glacial lakes, a key indicator of potential glacial lake outburst floods.[340] In permafrost research, ML performed analysis on the distribution of retrogressive thaw slumps[341] and extraction of ice-wedge polygons.[342] Beyond that, AI has been used to improve the quantification of sea ice surface coverage types, and also to extract Antarctic ice shelf fronts from Sentinel-1 Imagery[343] and to classify ice crystal habitats more precisely than traditional methods.[344] Therefore, AI plays a key role in promoting frontiers in cryospheric research by enabling the detection and characterization of specific features within cryospheric elements.

Properties retrieval. Different from traditional and complex physical models, AI enables simplified yet accurate property retrieval by modeling multivariate nonlinear relationships between cryospheric element parameters and image characteristics. This paradigm shift has led to significant advancements in understanding cryospheric processes. For example, AI improves the retrieval accuracy of the cryosphere properties in coalition with conventional algorithms, such as retrieval of snow depth[345] and estimates of snow water equivalent,[346,347] providing new insights to hydrological processes in cold regions. AI helped to solve the problem of detecting each internal ice layer uniquely to estimate their thickness accurately, thus providing crucial insights for assessing the contributions of ice sheets to sea level rise.[348] In permafrost research, AI has been applied to estimate mean annual ground temperature and active layer thickness and to estimate the thaw depth variations at seasonal scale.[349] AI has achieved better performance in Arctic sea ice thickness estimation, a key indicator of Arctic climate change.[332] In addition, AI has helped to reconstruct the winter glacier mass balance, a quantitative expression of glacier volume change through time, filling the gap in ground observations and providing valuable insights into long-term glacier volume changes.[350] Therefore, AI has led the way in streamlined and accurate cryosphere property retrieval.

Trend projection. AI significantly improves trend forecasting across diverse and complex conditions by developing sophisticated models that enhance spatiotemporal scope and precision. AI facilitates the investigation of historical cryospheric changes of possible trends, such as improving the prediction sensitivity of arsenic or manganese in groundwater and identifying trends that may not be apparent through traditional methods alone.[351] Similarly, AI was used to model the future responses of permafrost to climatic changes,[352] including permafrost degradation trends,[331] overcoming limitations of environmental conditions. In addition, AI was applied to estimate snow avalanche hazards for a better prediction of occurrence and magnitude.[353] AI also advanced the range of accurate sea ice forecast.[354] Regarding iceberg research, AI has been used to estimate the surface area and masses of icebergs,[334] which has operational difficulties in large-scale monitoring by observational and remote sensing methods. Therefore, AI-driven approaches significantly propel trend forecasting and predictive modeling within the cryosphere, providing valuable insights into historical changes, future projections, and operational challenges.

Cryospheric water storage dynamics and sea level rise. The cryosphere, a critical component of Earth's climatic system, is rapidly diminishing due to the effects of global warming. This trend is particularly evident in glaciers, including the massive Greenland and Antarctic ice sheets, which are experiencing accelerated mass loss. Moreover, sea ice coverage and snow extent are decreasing, while permafrost is undergoing significant degradation. This shrinking cryosphere is directly contributing to rising sea levels, posing imminent and long-term threats to low-lying coastal areas and small island nations. In addition, in mountainous regions and high plateaus, the reduction of cryospheric elements is causing fluctuations in river runoff, exacerbating water scarcity and increasing the risk of flooding in vulnerable areas. Cryospheric elements, such as glaciers, snowpacks, permafrost, sea ice, and ice caps, possess 3D or stereo characteristics. Traditional Earth observation methods often provide surface properties or limited-depth information, hindering comprehensive assessments of cryospheric elements. AI presents an opportunity to enhance our understanding of the 3D properties of cryospheric elements. For instance, AI can provide improved models of the active layer in permafrost and quantitatively assess the future conditions of permafrost.[349] AI-enhanced algorithms can better align with field data of snow depth.[345] Similarly, AI can improve sea ice thickness estimation algorithms to predict changes.[355] Utilizing AI for assessing mass balance from ice sheet volumes has the ability to estimate its contribution in sea-level rise, offering a new methodology of climate change studies.[356] In addition, AI has improved the precision of identifying each internal ice layer thickness in radar images, overcoming the limitations of traditional feature detection.[348] By combining Earth observation technologies, physical modeling, and AI techniques, researchers can delve deeper into the interior of the cryosphere, gaining crucial insights into its formation, evolution, and distribution. This integrated approach not only improves our understanding of cryospheric stereoscopic characteristics, but also enhances climate change research, particularly concerning cryosphere melting and its implications for sea-level rise.

Biosphere

Recent advances in satellites and aerial missions have led to the accumulation of ecological data streams, leading to the development trend of ML and DL models to advance our knowledge of the biosphere, including ecological parameter inversion and characteristics mapping.[113,357–362]

Vegetation properties mapping. Utilizing automatic learning of relationships between hundreds of bands and target variables, ML techniques such as decision trees, neural networks, and support vector machines have demonstrated exceptional efficiency in mapping vegetation structural and biochemical properties, encompassing leaf chlorophyll content, vegetation nitrogen, canopy cover, and leaf area index. In addition, ML algorithms play a crucial role in upscaling carbon fluxes (e.g., gross primary production, net ecosystem exchange, and ecosystem respiration) at regional and global scales.

Extracting vegetation variables is essential for evaluating how vegetation responds dynamically to fluctuating environmental conditions.[12] Utilizing automatic learning of relationships between hundreds of spectral bands and target variables, ML techniques such as decision trees, neural networks, and support vector machines have displayed outstanding performance in mapping vegetation structural and biochemical properties. These advanced algorithms effectively quantify parameters such as leaf chlorophyll content,[363,364] vegetation nitrogen,[365] canopy cover,[364,366,367] and leaf area index,[368–370] showcasing a substantial improvement over traditional empirical methods. These AI-driven models offer not only increased accuracy but also remarkable scalability and adaptability across different environmental conditions.[371–373]

Ecological parameter retrieval. In addition to mapping the vegetation properties and carbon fluxes, AI has advanced the precise identification of critical ecological parameters that were previously challenging to detect quickly and widely in terms of fine scale. Specifically, AI has achieved better performance in 3D structural parameters of forests such as leaf morphology,[91] tree height,[374] tree diameter at breast height,[375] and ground vegetation canopy size.[39] Similarly, AI techniques have also been employed in marine plankton structure.[376] In addition, AI helped to solve the problem of detecting and monitoring ecological disturbance.[370] Previous attempts have been based on laborious and complex handcrafted extraction of image features, but in recent years it has been shown that sophisticated convolutional neural networks can learn to extract relevant features automatically,[377] without human intervention. Automated image interpretation with convolutional neural networks performs very well for monitoring forest diseases and pests, close to human performance, and that makes professional field campaigns less costly.[225] In agricultural monitoring research, AI promotes the identification of malnourished crops, thereby assisting in the precise management of farmland.[378] Furthermore, AI has made significant progress in fine-scale geographic information simulation and prediction. Specifically, the rapid development of DL has notably enhanced the precision of urban characteristics simulating refined features more precisely than traditional methods.[379] AI also advanced the refined simulation of surface temperature and addresses the previously unresolved issue of fine simulation of extreme urban heat island effects.[380]

Fine-scale ecology analysis. On even finer scales, AI has achieved better performance in identifying ecological elements, promoting quantitative research on micro-ecosystems. In the research of diagnosing insects, AI techniques have reached 97% accuracy and outperformed a leading taxonomic expert.[381,382] For the identification and classification of vegetation pollen, DL technology has achieved automated pollen analysis methods,[383] which greatly solves the labor cost of labor-intensive pollen analysis in the past and significantly improves analysis efficiency. In addition, as a crucial means of extracting geographic information, classification technology has evolved further with the aid of AI foundations.[367,384] Currently, DL exhibits significant advantages in urban canopy detection[370] and tree species classification,[309] among others. By training with a large amount of data, DL-based models can achieve good prediction results for complex phenomena, such as crop element classification[385] and high-precision urban land element classification.[386] Simultaneously, existing experimental results demonstrate the superiority of the proposed AI model for both road detection and centerline extraction tasks.[387] Meanwhile, the integration of DL with high-resolution remote sensing images enables the refinement of ground feature statistics, which has advantages for separating and interpreting the desired information accurately over traditional remote sensing algorithms. For example, the U-Net neural network was employed to count trees in Africa,[377,388] which has operational difficulties in large-scale monitoring by observational and remote sensing methods. Overall, there is little doubt that there are many opportunities for trait-based ecology to benefit from the integration of computer vision and AI.

Global carbon budget. Accurate assessment of carbon dioxide uptakes and emissions of the terrestrial biosphere is critical to better understand the global carbon cycle, support the development of climate policies, and project future climate change.[93,384] AI plays an extremely important role in integrating satellite remote sensing and carbon fluxes from *in situ* observations to achieve high-precision, high-resolution scientific data on carbon fluxes of terrestrial ecosystems at regional and global scales.[228,389] For example, ML has been applied to estimated global plant gross primary production, net ecosystem exchange, ecosystem respiration, and soil respiration by integrating multi-source remote sensing data (i.e., various temperature, moisture, and plant production-related remote sensing products) and carbon fluxes data from ground observations.[390,391] The comparative advantages of AI over traditional methods are primarily due to its ability to effectively incorporate nonlinear relationships between remote sensing data and carbon fluxes. Thus, AI could assist the global carbon budget by providing more accurate and higher-resolution global plant production and ecosystem respiration detection.

Other domains

In addition to the aforementioned five spheres, AI is also significantly involved in other domains such as anthroposphere and inter-/cross-spheres, along with the engagement in sustainable development, opening new perspectives for analysis, interpretation, and fostering a more balanced relationship between human society and Earth's systems.[392]

Human activities understanding. AI plays a crucial role in comprehending and managing Earth's complex systems and environments, serving as a formidable toolset to glean insights, anticipate trends, and devise effective strategies for sustainable development and resource management. AI's multifaceted applications are particularly evident in its utilization by scientists for the analysis of real-time video streams derived from surveillance cameras and satellite imagery. This analytical prowess enables behavior analysis and large-scale monitoring of human activities, thereby offering invaluable insights into lifestyle patterns and social dynamics.[393] By harnessing AI-driven analytics, researchers can discern nuanced behavioral patterns, track movement trends, and identify emergent phenomena, facilitating a deeper understanding of human interactions with the environment and informing evidence-based decision-making processes.

Furthermore, AI serves as a cornerstone in the realm of urban development assessment, facilitating comprehensive analyses including diverse facets such as urban expansion, infrastructure changes, etc.[394] Leveraging AI-powered algorithms, urban planners and policymakers can assess the spatial dynamics of urban growth, anticipate infrastructure demands, optimize transportation networks, and devise sustainable land-use strategies. By amalgamating geospatial data with advanced analytical techniques, AI empowers stakeholders to make informed decisions aimed at fostering resilient, inclusive, and environmentally sustainable urban environments.

In tandem with its applications in physical environment monitoring, AI assumes a pivotal role in unraveling the intricacies of human behavior and preferences in the digital sphere. Social media analysis augmented by AI algorithms offers a potent lens through which online behavior and preferences can be discerned, thereby facilitating targeted advertising, personalized recommendations, and sentiment analysis.[146,395] By scrutinizing vast troves of user-generated content, AI-driven analytics can unveil latent trends, identify influencers, and gauge public sentiment, thereby enabling businesses and marketers to tailor their strategies to resonate with their target audience effectively.

In a word, AI's integration into Earth's complex systems and environments represents a paradigm shift in our ability to comprehend, monitor, and manage the multifaceted interplay between human activities and the natural world. By harnessing AI-driven analytics, researchers, policymakers, and businesses can unlock unprecedented insights, foster informed decision-making, and pave the way for a more sustainable and resilient future. However, it is imperative to acknowledge and address the ethical, privacy, and equity considerations inherent in the deployment of AI-powered systems, ensuring that these technologies are leveraged responsibly to serve the collective interests of humanity.

Spheres' interactions. AI has emerged as a powerful tool for capturing inter-layer relationships and enhancing simulations of biogeochemical cycles.[393] By leveraging AI techniques, such as DL, researchers can gain deeper insights into Earth's historical evolution and phenomena such as the snowball Earth event.[168] One notable advantage of AI in this context is its ability to improve computational efficiency[396] and parameter optimization,[397] thereby facilitating more accurate and robust simulations. In addition, AI aids in predicting matter exchange patterns and developing effective adaptation strategies to manage environmental changes.

Furthermore, AI contributes to refining our understanding of Earth's energy budget by integrating DL algorithms with remote sensing applications and incorporating biogeophysical feedback into models of the water cycle.[398,399] This interdisciplinary approach enables researchers to assess land surface changes and their impacts on energy budgets. Moreover, AI helps address the risks associated with over-parameterization in models, ensuring that simulations remain realistic and reliable. By identifying critical thresholds that trigger extreme events in Earth's systems, AI plays a crucial role in various applications, including volcano alerts,[400] groundwater mapping,[401] and studying climate-vegetation relationships.[402] This capability is crucial for improving early warning systems and mitigating the impacts of natural disasters on human populations and ecosystems.

The potential of AI extends beyond individual applications to regulating inter-layer dynamics and foreseeing thresholds that transform interactions at different scales. This proactive approach to exploring Earth's systems and managing its resources holds promise for sustainable Earth management. By leveraging AI technologies, researchers in geoscience can better anticipate and respond to environmental challenges, paving the way for more effective conservation efforts and informed policy decisions.

In conclusion, AI offers significant opportunities for advancing our understanding of Earth's complex systems and enhancing our ability to manage and protect the planet. By harnessing AI's capabilities in capturing inter-layer relationships, optimizing simulations, and identifying critical thresholds, researchers can contribute to proactive exploration and sustainable Earth management. However, realizing this potential requires continued interdisciplinary collaboration and the responsible deployment of AI technologies in geoscience research and environmental conservation efforts.

Sustainable development goals. The United Nations' 2030 Agenda outlines 17 interlinked goals that are set to solve development issues in economic, social, and environmental dimensions and realize sustainable development by 2030.[403] These goals interrelate closely with the Earth's spheres (lithosphere, hydrosphere, atmosphere, biosphere, and anthroposphere), aiming to ensure their equilibrium for human well-being and environmental sustainability. The appeal of leveraging AI to advance social benefits and achieve sustainable development goals (SDGs) has captured the attention of numerous practitioners and researchers.[404,405] For instance, in exploring the 169 targets outlined for the 17 goals, Vinuesa et al.[406] demonstrated that AI serves as an enabler for 134 targets while acting as an inhibitor for 59 targets. Gupta et al.[407] and Nasir et al.[408] delved into discussions about the implications of AI on the SDGs at the indicator level.

(1) *Economic sustainable development goals.* The technological benefits facilitated by AI also hold the potential to positively impact the attainment of several SDGs within the Economy group (SDGs 8, 9, 10, 11, and 12). Acemoglu and Restrepo indicate a net positive effect of AI-enabled technologies linked to increased productivity, highlighting potential negative consequences, particularly heightened inequalities.[409] If future markets heavily rely on data analysis and these resources are not equitably available in low- and middle-income countries, it could significantly widen the economic gap, exacerbating inequality even within nations.[410]

(2) *Social sustainable development goals.* For SDGs 1, 2, 3, 4, 5, 7, 16, and 17, in the social group, AI acts as an enabler for all the targets by supporting the provision of food, health, water, and energy services to the population, enhancing poverty mapping, identifying vulnerable populations, and optimizing resource allocation.[411,412] AI-based applications, including smart traffic management, waste management, and energy-efficient infrastructure, etc., contribute to developing sustainable and resilient urban developments.[413,414]

(3) *Environmental sustainable development goals.* The potential of AI extends to the analysis of extensive interconnected databases for collaborative initiatives aimed at environmental preservation (SDGs 6, 13, 14, and 15).[411] AI aids in water management through predictive analytics, monitoring water quality, and optimizing distribution networks.[415] AI is also poised to create low-carbon energy systems with the integration of renewable energy and essential components in climate 800 change, such as detecting the forest changes in satellite images to support habitat monitoring and decision-making.[416,417]

LARGE MODELS IN GEOSCIENCE

In this section, our principal objective is to elucidate the most recent developments associated with large models in geoscience,[418] alongside the presentation and summary of representative geoscience pre-trained foundation models.

Progress and application of large models in geoscience

The advent of large language models, prominently illustrated by ChatGPT, has significantly advanced diverse domains, concurrently empowering AI technologies to facilitate remarkable scientific progress, notably in geoscience. This is achieved through the autonomous calibration of billions of parameters during training, thereby enhancing representational capacity and learning capability.[68,419–422] The application of large models in geoscience, despite its unique challenges, has already demonstrated its huge revolutionary potential over traditional methods,[18,146,171,211,419,423–425] with the most noteworthy advances in the fields of remote sensing, atmosphere, ocean, and hydrology.[323,426–430]

Specifically, the remote sensing domain owns the most diverse data in the entire Earth science field.[431,432] General applications such as object detection, semantic segmentation, scene classification, and change detection from various data sources promoted the development of large models in remote sensing, such as the largest spectral remote sensing foundation model,[433] with an effective method for expanding and fine-tuning ViT.[434] Recently, AI Earth—based on a universal segmentation model (AIE-SEG) —was proposed by Alibaba to quickly extract any target in remote sensing images, achieving unified image segmentation tasks and rapid extraction of "zero samples of all things" without any labeled data. A new AI model called "segment anything model" from Meta AI can "cut out" any object in any image with zero-shot generalization to unfamiliar objects and images, without the need for additional training.[435] IBM and NASA have also teamed up to develop an open-source, geospatial foundation model that will enable researchers and scientists to utilize AI to track the amount of satellite data.[436] Furthermore, there is rapid development in multimodal remote sensing large models. For instance, SkySense[158] is a generic billion-scale model pre-trained on a curated multi-modal remote sensing imagery dataset with 21.5 million temporal sequences. In addition, large-scale vision-language models, such as EarthGPT,[437] have garnered significant attention in the remote sensing field, aiming to unify various remote sensing tasks and multi-sensor images. In a general sense, it can be observed that the utilization of large computer vision models and the efficient exploitation of vast remote sensing datasets to enhance the recognition of various targets represents a prominent trajectory in the evolution of large remote sensing models.

In the climate and weather domains, numerous large models with a great amount of data and parameters have been trained for predictions. For example, a Fourier forecasting neural network (FourCastNet) is proposed to provide immediate accurate short to medium-range global weather predictions.[438] The predictive outcomes derived from the FourCastNet model have been meticulously juxtaposed with the findings of the integrated forecasting system. It has been ascertained that the FourCastNet model exhibits substantial advantages across a multitude of performance indicators, with a particular emphasis on its notable progress in the domain of precipitation forecasting. Notably, the accuracy of the FourCastNet model surpasses that of other ones by an impressive margin, exceeding 20%. Pangu-Weather,[187] which harnesses the power of the 3D Earth-specific Transformer, has been empirically demonstrated to yield superior results, accompanied by a remarkable acceleration of 10,000 times, in contrast to the ECMWF. The proposal of NowcastNet,[232] a nonlinear nowcasting model for extreme precipitation, signifies a novel approach that unifies physical-evolution schemes and conditional-learning methods within a neural network framework. This model has proven its capacity to skillfully forecast extreme precipitation events characterized by advective or convective processes, previously deemed challenging to predict. MetNet-3,[439] a collaborative development by Google and DeepMind, has enhanced high-resolution predictions of several weather variables, encompassing precipitation, surface temperature, wind speed, and wind direction, for a forecast horizon extending up to 24 h. GenCast[440] proposes a generative model for global medium-range ensemble weather forecasting up to 15 days ahead, utilizing a diffusion model to sample ensembles from future weather trajectories. In addition, the swift advancement of large language models has positively impacted climate-related endeavors. For example, ClimateGPT[441] serves as a specialized conversational agent for climate change and sustainability topics in English and Arabic.

Concurrently, there have been recent propositions in the development of general geoscientific large-scale models. In the context of hydrology, a foundation platform, HydroPML,[323] is proposed for hydrological applications based on physics-aware ML. It bridges the gap between large language models and process-based hydrology, offering a range of applications, including but not limited to rainfall-runoff-inundation modeling,[122] real-time flood forecasting,[321] and cutting-edge methods to enhance water security and foster resilient water management. The first-ever large language model in the ocean domain, OceanGPT,[429] is introduced as an expert in various ocean science tasks. In the domain of disaster management and response, Disaster Response GPT is proposed to provide a versatile and adaptive framework for addressing various types of disasters and their associated challenges.[442] Furthermore, large models for time series forecasting, including variables such as wind and weather, have been proposed, leveraging a transformer backbone and zero-shot transfer.[443]

In summary, substantial advancements have been made in remote sensing and climate domains by deploying large models and effectively utilizing extensive datasets. However, widespread adoption of these methods on a broad scale remains challenging, particularly in extreme weather prediction. Progress in other geoscience areas, such as disaster prevention and hydrology, has been hindered by limited access to datasets and computing resources, slowing down the development of large language models. In the future, developing a unified, interpretable, and continuously learning large model to address the complexities and scales of geoscience will be a focus of ongoing exploration.

Pre-training of large geoscience models

Table 1 illustrates the schematic representation of the foundation of pre-trained models in geoscience. In the realm of remote sensing, various approaches have emerged, for instance, MoCo-V2 with geographic location serving as an agent task in conjunction with contrast learning for base model training,[154] CSPT using knowledge migration and image mask learning to enhance the expressive capability of the pre-trained model,[444] SeCo constructing positive-negative sample pairs from different seasons to effectively utilize unlabeled multi-seasonal data.[445] Wuhan University introduced the Billion Visual Transformer model,[446] exploiting a masking strategy for pre-processing, and achieved notable performance in image classification, target detection, and semantic segmentation. SatMAE,[447] proposed by Stanford University, adopts a grouped masking strategy for multi-temporal and multi-channel multispectral images. Recently, Hong et al.[433] designed the first and largest customized foundation model for spectral remote sensing data, i.e., SpectralGPT, achieving state-of-the-art performance in various downstream applications. Simultaneously, the work[448] combines SAR and multispectral images for a contrast learning approach. Another study[449] employs contrast learning, image filling, and deformation prediction as agent tasks to enhance the generalization of the pre-trained model. Researchers at the University of California, Berkeley focus on spatial scale information, modeling low-frequency and high-frequency details separately in the reconstruction layer.[450] In addition, Hong et al.[451] explored multimodal fusion on various image types, including optical images, SAR images, digital elevation models, and MAP data,[452] which innovated a new paradigm of multimodal AI big models for Earth observation, unlocking the Earth observation capability of remote sensing big data.[453] Presto reconstructs time series images through stacking and employing randomized masking strategies. Furthermore, GFM employs a teacher-student two-stream network on large-scale datasets,[454] excelling in scene classification, change detection, and semantic segmentation. Satvit explores the role of the MAE framework in analyzing satellite remote sensing data.[455]

In a distinct domain, ClimaX is pre-trained on the CMIP6 climate dataset,[428] offering versatility in weather and climate tasks. Notably, K2,[459] a 7 billion parameter Earth science language model from Shanghai Jiao Tong University, utilizes a two-stage construction involving pre-training on a high-quality Earth science corpus and instruction fine-tuning with a geosignal dataset. In contrast, general visual models such as Sky Eye and SenseEarth 3.0 improve remote sensing interpretation efficiency, leveraging Transformer-like backbones and self-supervised learning.

In summary, algorithms designed for processing remote sensing images exhibit variations in their emphasis on RGB, multispectral, or hyperspectral data, tailored for application to specific downstream tasks. Notably, contemporary climate and geoscience models such as K2 and ClimaX exemplify advancements in addressing challenges within these domains, showcasing enhanced efficiency and robustness for applications in Earth science. Despite the immense potential of large geoscience models, common research teams (usually small groups) encounter numerous impediments in embracing large-scale (pre-trained) models. Chiefly, constraints in resources, encompassing limited funding and manpower, impede their capacity to conduct research and development effectively. In addition, the intricacy of large-scale models poses a formidable learning curve for small teams, who may grapple with acquiring expertise across diverse disciplines such as ML and natural language processing. In the long term, the absence of access to comprehensive datasets and formidable competition from large tech companies further impede their progress. Legal and ethical considerations also present challenges, as small teams may lack the resources to adeptly navigate intricate issues such as privacy and accountability. Overall, surmounting these hurdles will necessitate strategic investments, collaboration, and concerted efforts to address legal and ethical concerns.

Deep-time digital Earth

Delving into the deep-time history of Earth is seen as a promising avenue to unravel the mechanisms of Earth's evolution, expose climate change patterns, identify natural resources, and envisage the future of our planet.[171,460] The advent of big data science in recent decades provides a valuable opportunity to tackle these questions. To expedite exploratory studies of Earth's evolution, there is a pressing need for an equitable, integrated database. To achieve this goal, the Deep Time Digital Earth (DDE) project is proposed as the inaugural "large-scale scientific project" by the International Union of Geological Sciences. This initiative aims to facilitate deep-time, data-driven discoveries through collaborative efforts across nations and disciplines.[461] Moreover, it introduces an open data platform to establish connections between existing deep-time geocounts and integrated geological data.

(1) *Earth's life evolution*. The synergy of AI and data science has significantly advanced our comprehension of Earth's life evolution, particularly concerning early complex life and mass extinctions. For instance,

Table 1. Representative pre-trained foundation models in geoscience

Application field	Model	Pre-trained model	Objectives
Remote sensing	CSPT[444]	ViT	improving the expressive ability of the pre-trained model
	RingMo[456]	ViT/Swin Transformer	a remote sensing foundation model with masked image modeling
	Scale-MAE[450]	Transformer	a pre-trained framework that introduces scale invariance into encoders that are used for a diverse set of downstream tasks
	SatMAE[447]	Transformer	pre-training Transformers for temporal and multi-spectral satellite imagery
	pre-trained ViT[427]	ViT	remote sensing foundation model
	GFM[454]	ViT	building geospatial foundation models via continual pre-training
	SatViT[455]	ViT	pre-training transformers for Earth observations
	Masked ViT[457]	ViT	self-supervised masked image reconstruction to advance transformer models for hyperspectral remote sensing imagery
	SpectralGPT[433]	ViT	the first customized foundation model designed explicitly for spectral remote sensing data
Weather and climate	Earthformer[458]	Transformer	a space-time Transformer for Earth system forecasting
	FourCastNet[438]	Fourier Neural Operator	provide accurate short- to medium-range global predictions
	GraphCast[157]	GNN	medium-range global weather forecasting
	NowcastNet[232]	physics-conditional generative network	a nonlinear nowcasting model for extreme precipitation
	MetNet[439]	U-Net + ViT	high-resolution predictions of several core weather variables
	Pangu-weather[187]	3D Transformer	accurate medium-range global weather forecasting
	ClimateX[428]	ViT	a foundation model for weather and climate
Others	K2[459]	Generative model (LLaMA-7B)	Earth science large language model
	DisasterResponseGPT[442]	Generative model	provide a versatile and adaptive framework for disasters
	OceanGPT[429]	Generative model	a large language model for ocean science tasks

ML methods are employed to analyze deep-time marine Paleozoic data, unraveling the impact of environmental changes on biodiversity.[462] The pulsed extinction of early complex life was further corroborated through network analysis of Ediacaran fossils.[463] Furthermore, the DDE project aims to integrate and interconnect existing deep paleontological and stratigraphic databases, leveraging DL and other AI tools to expedite biological data-driven discoveries.[464]

(2) *Earth's material evolution.* In the context of Earth's material evolution, current AI-driven approaches strive to propel the evolution and discovery of minerals, rocks, sediments, and fluids. Noteworthy examples encompass the evolution of minerals,[465] the cycling of sediments,[466] and the interpretation of plate tectonics.[467] In addition, AI-driven discovery necessitates the integration of existing geomaterial databases by the DDE, enhancing spatial and temporal coverage as well as resolution in the discovery of geomaterials.

(3) *Geography's evolution.* Geography's evolution holds paramount significance in various domains, including mineral and energy resource assessment, Earth hazard prediction, comprehending Earth's history, and forecasting the future. The correlation of deep Earth science databases with paleogeographic reconstruction databases is an important goal of DDE. Supported by big data analysis techniques, this combination has been widely used in the field of paleontology,[468] paleoclimatology,[469] and geodynamics.[470]

(4) *Paleoclimate's evolution.* The exploration of paleoclimate assumes a crucial role in understanding the interaction between Earth and life in producing climate extremes and forecasting future climate changes.[471] AI's strengths in data processing, hypothesizing, and predicting within Earth science research substantially facilitate paleoclimate reconstruction.[472] Assisted by AI, the DDE can reconstruct the history of paleoclimate and paleoatmosphere, relying on various minerals, rocks, and geochemical indicators preserved in Earth material.[473]

In summary, the establishment of a unified representation model to head the construction of an integrated Earth science knowledge map is one of the key programs of DDE,[474,475] and a series of knowledge graphs have emerged, such as the paleoclimate knowledge graph,[476] standard carbonate microfacies,[477] and academic knowledge graph.[478] With the continued emergence of geoscientific macrolanguage models (such as K2[459]), AI has dramatically changed the traditional paradigm of geoscientific research. By harmonizing and integrating deep Earth data, geological knowledge, and advanced techniques in data science and AI, DDE is poised to advance solutions for the significant challenges in Earth evolution research, understanding the past, present, and future of our planet.

CHALLENGES AND OUTLOOKS IN AI FOR GEOSCIENCE

The numerous cases and advanced techniques outlined in the previous sections solidly prove that AI is an expert technology at deciphering complex relationships in the Earth system and predicting environmental responses with unprecedented accuracy. However, this is not the end of the journey; there remain ongoing challenges and opportunities in the field of research. This section poses the challenges and future perspectives to promote the co-development of AI and geoscience.

Unsolved challenges of AI for geoscience

There are many unsolved challenges in AI for geoscience, particularly at the intersection of these two fields. These challenges arise from interdisciplinary complexities, making it difficult for scientists to identify and address the problems.

Ethical considerations play a crucial role across all stages of geoscience disciplines, encompassing data collection, analysis, and distribution.[479] High-resolution data, for example, raise privacy concerns,[480] while socio-economic analyses can lead to stigmatization if not handled carefully.[481] The demand for explainability grows as AI applications extend their reach into policy-making, requiring models to be both transparent and justifiable.[146] Addressing these ethical challenges involves adhering to robust ethical frameworks and guidelines, promoting a culture of geoethical thinking and social responsibility among researchers.

Moreover, due to the biased learning knowledge by AI, the adeptness of AI in modeling complex relationships brings about vulnerabilities related to data security.[482,483] The potential for data bias and tampering poses significant risks, potentially leading to misrepresentations of geographical features and misguided policy decisions. To mitigate these risks, a multifaceted approach, including robust data validation and enhancements in AI learning specifications, is essential. These strategies not only fortify data integrity but also improve the resilience of AI systems against malicious manipulations.

Despite the exceptional capabilities of AI, the demand for computing resources and the costs associated with data acquisition and processing present substantial challenges.[17] The computational intensity required for models, such as predicting global climate[187] or global forest fire interactions,[484] necessitates substantial investment in computational and memory resources, often beyond the reach of many geoscientists. Moreover, the AI models should be energy efficient so that they can also contribute to the NetZero agenda. To optimize performance and reduce expenses, strategies such as leveraging cloud computing, applying transfer learning, and enhancing data management practices are vital.[284] These approaches help in managing the high costs and logistical demands of extensive data processing, ensuring that AI applications remain both viable and effective.

Emerging challenges in new paradigm of hybrid models

Hybrid models, leveraging the strength of physics-based models and AI, are starting to show their charming potential as a new research paradigm in geoscience. Despite their potential, they present challenges in the development of the paradigm.

The first challenge is the uncertain interpretability within the model. While the structure of hybrid models seems to maintain physical plausibility, and the AI component can even effectively compensate for structural deficiency in physics-based counterpart,[485,486] there remains a critical concern. Often, the balance between physics-based and AI components in hybrid models may be overlooked due to a lack of integration knowledge within the "gray box." The work by Acuña Espinoza et al.[487] suggests that AI-based parameterization may learn incorrect behaviors and overwrite the physical interpretability in the hybrid hydrological models, despite enhancing performance. This compensatory capability of AI raises questions about the true hydrological interpretability of outputs from hybrid models. It also calls for a more cautious use of hybrid models in geoscience applications, particularly when the primary objective is to decipher geophysical processes rather than merely improve prediction accuracy.

Another challenge in advancing this paradigm is extending these hybrid models to accommodate large datasets and complex system interactions inherent in global geoscience applications. As these models scale, the structural deficiencies in the physics-based part of the hybrid model will be magnified,[485] and maintaining a balance between AI fitting capabilities and physical interpretability will become increasingly difficult. Therefore, large models currently applied in geoscience, such as the FourCastNet and Pangu-Weather models, are still predominantly in the data-driven paradigm and risk losing physical plausibility. This scaling issue highlights the need for a deep understanding of geophysical processes in hybrid models at the regional scale.

Outlook on AI for inter-spheres

While the application and knowledge of AI for intra-spheres are relatively comprehensive, exploring inter-spheres connection in geoscience reveals significant knowledge gaps.[488] These gaps arise from the challenges of integrating fragmented knowledge across disciplines when enhancing Earth system models. The complexity of cross-system dynamics and feedback mechanisms complicates the encoding of multidisciplinary and multi-domain knowledge. For instance, the biochemical and biophysical processes within the hydrological cycle[489] and the atmospheric-ocean interaction[490] are crucial cases for understanding the hydrological cycle and predicting phenomena such as the Madden-Julian Oscillation and El Niño Southern Oscillation, respectively. Yet, they exhibit gaps in multidisciplinary integration.

Undoubtedly, AI has demonstrated the potential to bridge these interdisciplinary gaps, as demonstrated by its successful application within individual domains. Several studies have already started to apply AI to forge connections across multiple spheres. For example, AI-powered prediction models have been used to forecast hurricanes by analyzing the complex interplay between ocean temperatures, atmospheric conditions, and land surface characteristics.[491] However, advancing AI development in the inter-sphere's context requires greater efforts, including more robust exchanges of expert knowledge and domain-specific insights.

Outlook of AI for exploring exoplanets

The lack of terrestrial data with viable and varied observational evidence represents a significant bottleneck in the development of geoscience. Terrestrial exoplanets, sharing similar geophysical processes, can complement the data gap. Planetary scientists suggest that the understanding of the cooling and transfer of heat from the interiors of terrestrial planets can help explain the geological evolution of Earth.[492] Furthermore, studying tidal interaction on low-mass planets can aid in understanding atmospheric circulation and meteorological phenomena on Earth.[493] This highlights the potential of exoplanet exploration to offer new insights into our own planet.

In contrast to knowledge transfer from exoplanets to Earth, there remain plenty of unknowns about the environment of exoplanets, frequently resulting in a less sophisticated understanding of their geophysical processes compared with Earth. Generally, discussions about exoplanet characteristics often simply rely on the knowledge of an exoplanet's mass, radius, or orbital distance. In this context, the power of AI can be used to decipher the high complexity of an exoplanet's system. Some works[494,495] suggest that AI approaches trained by biosignatures on Earth could be adapted for searching for life on terrestrial exoplanets. Interdisciplinary application of AI in geoscience, transferring from Earth to exoplanets, could enhance our understanding of these distant worlds' geophysical processes, thereby offering a fresh perspective on Earth in the future.

Future development of AI for geoscience

Our review demonstrates the necessity of advancing AI for geoscience research. Looking ahead, AI is poised to significantly enhance geoscience projects, supported by various government and authoritative endorsements. For example, the China Ministry of Science and Technology highlights AI as a pivotal tool for groundbreaking research across four strategic frontiers: deep space, deep sea, deep Earth, and "deep blue." Similarly, NASA regards AI as an essential component for future Earth explorations.[496]

Conversely, our review also acknowledges the profound and dynamic impact of AI on our understanding of geoscience and on decision-making processes. However, there is limited consensus on the regulations governing AI development and usage. The United Nations Educational, Scientific and Cultural Organization[497] and the European Union's General Data Protection Regulation[498] underscore the importance of ethical considerations, such as privacy, interpretability, and security in AI applications, which indicates the need for a model-data-driven paradigm to enhance transparency in research.

CONCLUSION

The research paradigms in geoscience started with physics-based models, followed by data-driven approaches, and merged into hybrid models. This review strives to delineate these paradigms, emphasizing the unexplored frontiers where cutting-edge AI techniques intersect with geoscience. We put a special focus on hybrid models, which, leveraging domain knowledge to guide AI models, often require less training data while maintaining comparable accuracy, thus offering enhanced efficiency and performance. The potential of large-scale AI models in geoscience is vast, yet its realization faces challenges unique to the domain, impeding its widespread adoption and implementation. The dichotomy between these paradigms—space centered on explicit adherence to physical rules versus the extraction of insights from immense data volumes—underscores the need for a balanced approach in contemporary geoscience.

In essence, the quest to comprehend Earth's intricacies demands an amalgamation of diverse methodologies and approaches. The synergy between traditional principles and modern AI-driven techniques holds immense promise, yet it also presents a spectrum of challenges that require concerted efforts to overcome. As geoscientists navigate this dynamic terrain, a harmonized blend of methodologies stands poised to unlock profound insights into our planet's mysteries, shaping the trajectory of geoscience in the years to come.

REFERENCES

1. Super, J. (2023). Geoscientists excluded. Nat. Geosci. ***16***(3): 194. https://doi.org/10.1038/s41561-023-01152-z.
2. Gu, B., van Grinsven, H.J., Lam, S.K., et al. (2021). A Credit System to Solve Agricultural Nitrogen Pollution. Innovation ***2***(1): 100079. https://doi.org/10.1016/j.xinn.2021.100079.
3. Wetherill, G.W. (1990). Formation of the Earth. Annu. Rev. Earth Planet Sci. ***18***(1): 205–256. https://doi.org/10.1146/annurev.ea.18.050190.001225.
4. Zimmer, C. (2005). How and Where Did Life on Earth Arise? Science ***309***(5731): 89. https://doi.org/10.1126/science.309.5731.89.
5. Marty, B.A., Zimmermann, L., Pujol, M., et al. (2013). Nitrogen Isotopic Composition and Density of the Archean Atmosphere. Science ***342***(6154): 101–104. https://doi.org/10.1126/science.1240971.
6. Elkins-Tanton, L.T. (2013). Evolutionary dichotomy for rocky planets. Nature ***497***(7451): 570–572. https://doi.org/10.1038/497570a.
7. Freed, A.M. (2012). Casting stress shadows. Nat. Geosci. ***5***(6): 371–372.
8. Cheng, H., Li, H., Sha, L., et al. (2022). Milankovitch theory and monsoon. Innovation ***3***(6): 100338. https://doi.org/10.1016/j.xinn.2022.100338.
9. Wang, P. (2021). Low-latitude forcing, A new insight into paleo-climate changes. Innovation ***2***(3): 100145. https://doi.org/10.1016/j.xinn.2021.100145.
10. Bergen, K.J., Johnson, P.A., de Hoop, M.V., et al. (2019). Machine learning for data-driven discovery in solid Earth geoscience. Science ***363***(6433): eaau0323. https://doi.org/10.1126/science.aau0323.
11. Ge, Y., Zhang, X., Atkinson, P.M., et al. (2022). Geoscience-aware deep learning, A new paradigm for remote sensing. Sci. Remote Sens. ***5***: 100047. https://doi.org/10.1016/j.srs.2022.100047.
12. Kumar, P., Debele, S.E., Khalili, S., et al. (2024). Urban heat mitigation by green and blue infrastructure, Drivers, effectiveness, and future needs. Innovation ***5***(2): 100588. https://doi.org/10.1016/j.xinn.2024.100588.
13. Cheng, H. (2020). Future Earth and Sustainable Developments. Innovation ***1***(3): 100055. https://doi.org/10.1016/j.xinn.2020.100055.
14. Tibi, R., Wiens, D.A., and Inoue, H. (2003). Remote triggering of deep earthquakes in the 2002 Tonga se-quences. Nature ***424***(6951): 921–925. https://doi.org/10.1038/nature01903.
15. Donner, R.V., and Barbosa, S.M. (2008). Nonlinear Time Series Analysis in the Geosciences. Lect. Notes Earth Sci. ***112***(2008): 37. https://doi.org/10.1007/978-3-540-78938-3.
16. Sharma, A.S., Baker, D.N., Bhattacharyya, A., et al. (2012). Complexity and Extreme Events in Geosciences, An Overview. Geophys. Monogr. Ser. ***196***: 1–16. https://doi.org/10.1029/2012GM001233.
17. Reichstein, M., Camps-Valls, G., Stevens, B., et al. (2019). Deep learning and process understanding for data-driven earth system science. Nature ***566***: 195–204. https://doi.org/10.5194/egusphere-egu24-15874.
18. Xu, Y., Liu, X., Cao, X., et al. (2021). Artificial intelligence, A powerful paradigm for scientific research. Innovation ***2***(4): 100179. https://doi.org/10.1109/iiccit55816.2022.10010688.
19. Wang, L., Zuo, B., Le, Y., et al. (2023). Penetrating remote sensing, Next-generation remote sensing for transparent earth. Innovation ***4***(6): 100519. https://doi.org/10.1016/j.xinn.2023.100519.
20. Sun, J., Wang, Y., Piao, S., et al. (2022). Toward a sustainable grassland ecosystem worldwide. Innovation ***3***(4): 100265. https://doi.org/10.1016/j.xinn.2022.100265.
21. Li, Y., Zhang, H., Huang, P., et al. (2024). Demonstration of 10 Gbps satellite-to-ground laser communications in engineering. Innovation ***5***(1): 100557. https://doi.org/10.1016/j.xinn.2023.100557.
22. Xu, Y., Wang, F., An, Z., et al. (2023). Artificial intelligence for science—bridging data to wisdom. Innovation ***4***(6): 100525. https://doi.org/10.1016/j.xinn.2023.100525.
23. Richard, S.M. (2006). Geoscience concept models. In Geoinformatics, Data to Knowledge (Geological Society of America). https://doi.org/10.1130/2006.2397(07.
24. Bokulich, A., and Oreskes, N. (2017). Models in Geosciences (Springer Handb. Model-Based Sci). https://doi.org/10.1007/978-3-319-30526-4_41.
25. Bruno, M.S. (2005). Physical Models. Encycl. Earth Sci. Ser. 769–771. https://doi.org/10.1007/1-4020-3880-1_245.
26. Oreskes, N. (2020). From Scaling to Simulation, Changing Meanings and Ambitions of Models in Geology (SwL, MSCEN), pp. 93–124. https://doi.org/10.1515/9780822390244-006.
27. Fowler, A. (2011). Mathematical Modelling. Math. Geosci. 1–63. https://doi.org/10.1007/978-0-85729-721-1.
28. Gerya, T. (2019). Introduction to Numerical Geodynamic Modelling (Cambridge University Press). https://doi.org/10.1017/9781316534243.
29. Winsberg, E. (2015). In Computer Simulations in Science, E.N. Zalta and U. Nodelman, eds. (Stanford University). https://doi.org/10.1007/978-3-031-38647-3_2.
30. Kimura, R. (2002). Numerical weather prediction. J. Wind Eng. Ind. Aerodyn. ***90***(12): 1403–1414.
31. Du, M., Peng, X., Zhang, H., et al. (2021). Geology, environment, and life in the deepest part of the world's oceans. Innovation. ***2***(2). https://doi.org/10.1016/j.xinn.2021.100109.
32. Berliner, L.M. (2003). Physical-statistical modeling in geophysics. J. Geophys. Res. ***108***(24). https://doi.org/10.1029/2002JD002865.
33. Green, D. (2014). Modelling Geomorphic Systems: Scaled Physical Models. Geomorphol.
34. Ehrendorfer, M. (1997). Predicting the uncertainty of numerical weather forecasts: a review. metz ***6***: 147–183.
35. Wang, W., Tschauner, O., Huang, S., et al. (2021). Coupled deep-mantle carbon-water cycle: Evidence from lower-mantle diamonds. Innovation ***2***(2): 100117. https://doi.org/10.1016/j.xinn.2021.100117.
36. Jiang, C., Feng, X., Guo, Y., et al. (2022). Data-driven modeling of solar coronal magnetic field evolution and eruptions. Innovation ***3***(3): 100236. https://doi.org/10.1016/j.xinn.2022.100236.
37. Lary, D.J., Alavi, A.H., Gandomi, A.H., et al. (2016). Machine learning in geosciences and remote sensing. Geosci. Front. ***7***(1): 3–10. https://doi.org/10.1016/j.gsf.2015.07.003.
38. Ouadfeul, S.-A., Jawak, S.D., Shirzadi, A., et al. (2023). Editorial: Artificial intelligence and machine learning in Earth science. Front. Earth Sci. ***10***. https://doi.org/10.3389/feart.2022.1090016.
39. Yang, X., Wu, J., Chen, X., et al. (2021). A comprehensive framework for seasonal controls of leaf abscission and productivity in evergreen broadleaved tropical and subtropical forests. Innovation ***2***(4): 100154. https://doi.org/10.1016/j.xinn.2021.100154.
40. Zhang, Z., Ni, W., Quegan, S., et al. (2024). Deforestation in Latin America in the 2000s predominantly occurred outside of typical mature forests. Innovation ***5***(3): 100610. https://doi.org/10.1016/j.xinn.2024.100610.
41. Krige, D.G. (1951). A Statistical Approach to Some Mine Valuation and Allied Problems on the Witwatersrand, D.G. Krige, ed. (Univ. of the Witwatersrand).
42. Pinto, J.A., Kumar, P., Alonso, M.F., et al. (2020). Kriging method application and traffic behavior profiles from local radar network database: A proposal to support traffic solutions and air pollution control strategies. Sustain. Cities Soc. ***56***: 102062. https://doi.org/10.1016/j.scs.2020.102062.
43. Shukla, K., Kumar, P., Mann, G.S., et al. (2020). Mapping spatial distribution of particulate matter using kriging and inverse distance weighting at supersites of megacity Delhi. Sustain. Cities Soc. ***54***: 101997. https://doi.org/10.1016/j.scs.2019.101997.
44. Mjolsness, E., and DeCoste, D. (2001). Machine Learning for Science: State of the Art and Future Prospects. Science ***293***(5537): 2051–2055. https://doi.org/10.1126/science.293.5537.2051.
45. Jansson, N.F., Allen, R.L., Skogsmo, G., et al. (2022). Principal component analysis and K-means clustering as tools during exploration for Zn skarn deposits and industrial carbonates, Sala area, Sweden. J. Geochem. Explor. ***233***: 106909. https://doi.org/10.1016/j.gexplo.2021.106909.
46. Li, J., Bioucas-Dias, J.M., and Plaza, A. (2012). Spectral-Spatial Hyperspectral Image Segmentation Using Subspace Multinomial Logistic Regression and Markov Random Fields. IEEE Trans. Geosci. Remote Sens. ***50***(3): 809–823. https://doi.org/10.1109/TGRS.2011.2162649.
47. Sarailidis, G., Wagener, T., and Pianosi, F. (2022). Integrating scientific knowledge into machine learning using interactive decision trees. Comput. Geosci. ***170***: 105248. https://doi.org/10.1016/j.cageo.2022.105248.
48. Liu, C., He, L., Li, Z., et al. (2018). Feature-Driven Active Learning for Hyperspectral Image Classification. IEEE Trans. Geosci. Remote Sens. ***56***(1): 341–354. https://doi.org/10.1109/TGRS.2017.2747862.
49. Lary, D.J. (2010). Artificial Intelligence in Geoscience and Remote Sensing. Geosci. Remote Sens. New Achiev. ***7***. https://doi.org/10.5772/9104.
50. Zhang, Z., Li, C., Wang, W., et al. (2024). Towards full-stack deep learning-empowered data processing pipeline for synchrotron tomography experiments. Innovation ***5***(1): 100539.
51. Zhu, X.X., Tuia, D., Mou, L., et al. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. IEEE Geosci. Remote Sens. Mag. ***5***(4): 8–36. https://doi.org/10.1109/MGRS.2017.2762307.
52. Galvão, S.L.J., Matos, J.C.O., Kitagawa, Y.K.L., et al. (2022). Particulate matter forecasting using different deep neural network topologies and wavelets for feature augmentation. Atmosphere-Basel ***13***(9): 1451. https://doi.org/10.3390/atmos13091451.

53. Baseer, M.A., Almunif, A., Alsaduni, I., et al. (2024). An intelligent optimized deep network-based predictive system for wind power plant application. Electr. Eng. 1–13. https://doi.org/10.1007/s00202-024-02377-w.
54. Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. NeurIPS ***25***. https://doi.org/10.1145/3065386.
55. Mou, L., Ghamisi, P., and Zhu, X.X. (2018). Unsupervised Spectral-Spatial Feature Learning via Deep Residual Conv-Deconv Network for Hyperspectral Image Classification. IEEE Trans. Geosci. Remote Sens. ***56***(1): 391–406. https://doi.org/10.1109/TGRS.2017.2748160.
56. Deng, P., Xu, K., and Huang, H. (2022). When CNNs Meet Vision Transformer: A Joint Framework for Remote Sensing Scene Classification. IEEE Geosci. Remote Sens. Lett. ***19***: 1–5. https://doi.org/10.1109/LGRS.2021.3109061.
57. Behrens, T., Schmidt, K., MacMillan, R.A., et al. (2018). Multi-scale digital soil mapping with deep learning. Sci. Rep. ***8***(1): 15244. https://doi.org/10.1038/s41598-018-33516-6.
58. Mou, L., Bruzzone, L., and Zhu, X.X. (2019). Learning Spectral-Spatial-Temporal Features via a Recurrent Convolutional Neural Network for Change Detection in Multispectral Imagery. IEEE Trans. Geosci. Remote Sens. ***57***(2): 924–935. https://doi.org/10.1109/TGRS.2018.2863224.
59. Karpatne, A., Ebert-Uphoff, I., Ravela, S., et al. (2019). Machine Learning for the Geosciences: Challenges and Opportunities. IEEE Trans. Knowl. Data Eng. ***31***(8): 1544–1554. https://doi.org/10.1109/TKDE.2018.2861006.
60. Yu, S., and Ma, J. (2021). Deep Learning for Geophysics: Current and Future Trends. Rev. Geophys. ***59***(3): e2021RG000742. https://doi.org/10.1029/2021RG000742.
61. Zhang, W., Gu, X., Tang, L., et al. (2022). Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: Comprehensive review and future challenge. Gondwana Res. ***109***: 1–17. https://doi.org/10.1016/j.gr.2022.03.015.
62. Maskey, M., Alemohammad, H., Murphy, K., et al. (2020). Advancing AI for Earth science: A data systems perspective. ESA EO Phiweek. https://doi.org/10.1029/2020eo151245.
63. Shen, C., Appling, A.P., Gentine, P., et al. (2023). Differentiable modelling to unify machine learning and physical models for geosciences. Nat. Rev. Earth Environ. ***4***(8): 552–567. https://doi.org/10.1038/s43017-023-00450-9.
64. Liu, C., Li, J., He, L., et al. (2021). Naive Gabor Networks for Hyperspectral Image Classification. IEEE Trans. Neural Netw. Learn. Syst. ***32***(1): 376–390. https://doi.org/10.1109/TNNLS.2020.2978760.
65. Jiang, S., Zheng, Y., and Solomatine, D. (2020). Improving AI System Awareness of Geoscience Knowledge: Symbiotic Integration of Physical Approaches and Deep Learning. Geophys. Res. Lett. ***47***(13): e2020GL088229. https://doi.org/10.1029/2020GL088229.
66. Willard, J., Jia, X., Xu, S., et al. (2020). Integrating physics-based modeling with machine learning: A survey. Preprint at arXiv. https://doi.org/10.1145/1122445.1122456.
67. Camps-Valls, G., Reichstein, M., Zhu, X.X., et al. (2020). Advancing deep learning for Earth sciences: From hybrid modeling to interpretability. IGARSS ***2020***: 3979–3982. https://doi.org/10.1109/IGARSS39084.2020.9323558.
68. Zhu, J.-J., Jiang, J., Yang, M., et al. (2023). ChatGPT and environmental research. Environ. Sci. Technol. ***57***(46): 17667–17670. https://doi.org/10.1021/acs.est.3c01818.
69. Wang, Y., Ait Ali Braham, N., Xiong, Z., et al. (2023). SSL4EOS12: A large-scale multimodal, multitemporal dataset for self-supervised learning in Earth observation. IEEE Geosci. Remote Sens. Mag. ***11***(3): 98–106. https://doi.org/10.1109/MGRS.2023.3281651.
70. Dramsch, J.S. (2020). Chapter One - 70 years of machine learning in geoscience in review. Adv. Geophys. ***61***: 1–55. https://doi.org/10.1016/bs.agph.2020.08.002.
71. Wang, Y., Albrecht, C.M., Ait Ali Braham, N.A.A., et al. (2022). Self-supervised learning in remote sensing: A review. IEEE Geosci. Remote Sens. Mag. ***10***(4): 213–247. https://doi.org/10.1109/MGRS.2022.3198244.
72. Ji, J., Wang, S., Li, H., et al. (2022). CHES: An astrometry mission searching for nearby habitable planets. Innovation. ***3***(4). https://doi.org/10.1016/j.xinn.2022.100270.
73. Ge, J., Zhang, H., Deng, H., et al. (2022). The ET mission to search for Earth 2.0s. Innovation ***3***(4): 100271. https://doi.org/10.1016/j.xinn.2022.100271.
74. Le, H., Rong, Z., and Wei, Y. (2023). Exploring the universe and protecting the Earth: Young Chinese scientists in action. Innovation ***4***(4): 100466. https://doi.org/10.1016/j.xinn.2023.100466.
75. Zhang, X. (2023). JWST's eyes on an alien world. Innovation ***4***(3): 100428. https://doi.org/10.1016/j.xinn.2023.100428.
76. Zheng, Y.C. (2020). Mars Exploration in 2020. Innovation ***1***(2): 100036. https://doi.org/10.1016/j.xinn.2020.100036.
77. Artemis - NASA. https://www.nasa.gov/humans-in-space/artemis.
78. China's Lunar and Deep Space Exploration. http://www.clep.org.cn/n487137/index.html
79. Yang, W., and Lin, Y. (2021). New Lunar Samples Returned by Chang'e-5: Opportunities for New Discoveries and International Collaboration. Innovation ***2***(1): 100070. https://doi.org/10.1016/j.xinn.2020.100070.
80. BepiColombo. https://www.esa.int/Science_Exploration/Space_Science/BepiColombo.
81. Yan, H., Li, H., Wang, S., et al. (2022). Overview of the LAMOST survey in the first decade. Innovation ***3***(2): 100224. https://doi.org/10.1016/j.xinn.2022.100224.
82. Li, H. (2022). Go beyond Hubble and go deeper in the universe. Innovation ***3***(5): 100305. https://doi.org/10.1016/j.xinn.2022.100305.
83. Sun, Z., Sandoval, L., Crystal-Ornelas, R., et al. (2022). A review of Earth Artificial Intelligence. Comput. Geosci. ***159***: 105034. https://doi.org/10.1016/j.cageo.2022.105034.
84. Steffen, W., Richardson, K., Rockström, J., et al. (2020). The emergence and evolution of earth system science. Nat. Rev. Earth Environ. ***1***(1): 54–63. https://doi.org/10.1038/s43017-019-0005-6.
85. Schellnhuber, H.J. (1999). Earth system analysis and the second Copernican revolution. Nature ***402***(6761): 19–23. https://doi.org/10.1038/35011515.
86. Tuia, D., Roscher, R., Wegner, J.D., et al. (2021). Toward a collective agenda on AI for Earth science data analysis. IEEE Geosci. Remote Sens. Mag. ***9***(2): 88–104. https://doi.org/10.1109/MGRS.2020.3043504.
87. Lovelock, J.E., and Margulis, L. (1974). Atmospheric homeostasis by and for the biosphere: The Gaia hypothesis. Tellus ***26***(1-2): 2–10. https://doi.org/10.3402/tellusa.v26i1-2.9731.
88. Li, K., Zheng, F., Cheng, L., et al. (2023). Record-breaking global temperature and crises with strong El Niño in 2023-2024. Innovation Geosci. ***1***(2): 100030. https://doi.org/10.59717/j.xinn-geo.2023.100030.
89. Hurrell, J.W., Holland, M.M., Gent, P.R., et al. (2013). The Community Earth System Model: A framework for collaborative research. Bull. Am. Meteorol. Soc. ***94***(9): 1339–1360. https://doi.org/10.1175/BAMS-D-12-00121.1.
90. Jensen, J.R. (2009). Remote sensing of the environment: An earth resource perspective. Cartogr. Geogr. Inf. Sci. ***27***(4): 311. https://doi.org/10.1559/1523040200.
91. Wang, F., Harindintwali, J.D., Wei, K., et al. (2023). Climate change: Strategies for mitigation and adaptation. Innovation Geosci. ***1***(1): 100015–100061. https://doi.org/10.59717/j.xinn-geo.2023.100015.
92. Suzuki, D. (2022). The Sacred Balance: Rediscovering Our Place in Nature (Greystone Books Ltd). https://doi.org/10.2307/1522232.
93. Chen, M., Lv, G., Zhou, C., et al. (2021). Geographic modeling and simulation systems for geographic research in the new era: Some thoughts on their development and construction. Sci. China Earth Sci. ***64***(8): 1207–1223. https://doi.org/10.1007/s11430-020-9759-0.
94. Claussen, M., Mysak, L., Weaver, A., et al. (2002). Earth system models of intermediate complexity: Closing the gap in the spectrum of climate system models. Clim. Dyn. ***18***: 579–586. https://doi.org/10.1007/s00382-001-0200-1.
95. Zhu, Z., Chen, M., Qian, Z., et al. (2023). Documentation strategy for facilitating the reproducibility of geo-simulation experiments. Environ. Model. Softw. ***163***: 105687. https://doi.org/10.1016/j.envsoft.2023.105687.
96. Beven, K., and Freer, J. (2001). A dynamic TOPMODEL. Hydrol. Process. ***15***(10): 1993–2011. https://doi.org/10.1002/hyp.252.
97. Arnold, J. (1994). SWAT—Soil and Water Assessment Tool. https://doi.org/10.1007/springerreference_62887.
98. Gironás, J., Roesner, L.A., Rossman, L.A., et al. (2010). A new applications manual for the Storm Water Management Model (SWMM). Environ. Modell. Softw. ***25***(6): 813–814. https://doi.org/10.1016/j.envsoft.2009.11.009.
99. Chen, C., Beardsley, R.C., Cowles, G., et al. (2012). An Unstructured-Grid, Finite-Volume Community Ocean Model: FVCOM User Manual (Massachusetts Institute of Technology Cambridge). https://doi.org/10.23919/oceans.2009.5422441.
100. Done, J., Davis, C.A., and Weisman, M. (2004). The next generation of NWP: Explicit forecasts of convection using the Weather Research and Forecasting (WRF) model. Atmos. Sci. Lett. ***5***(6): 110–117. https://doi.org/10.1002/asl.72.
101. Binkowski, F.S., and Roselle, S.J. (2003). Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component 1. Model description. J. Geophys. Res. ***108***(6). https://doi.org/10.1029/2001JD001409.
102. Guenther, A., Karl, T., Harley, P., et al. (2006). Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). Atmos. Chem. Phys. ***6***(11): 3181–3210. https://doi.org/10.5194/acpd-6-107-2006.
103. Keenan, T., Serra, J.M., Lloret, F., et al. (2011). Predicting the future of forests in the Mediterranean under climate change, with niche-and process-based models: CO2 matters. Glob. Chang. Biol. ***17***(1): 565–579. https://doi.org/10.1111/j.1365-2486.2010.02254.x.
104. Van Vliet, J., Bregt, A.K., Brown, D.G., et al. (2016). A review of current calibration and validation practices in land-change modeling. Environ. Model. Softw. ***82***: 174–182. https://doi.org/10.1016/j.envsoft.2016.04.017.
105. Jakeman, A.J., Letcher, R.A., and Norton, J.P. (2006). Ten iterative steps in development and evaluation of environmental models. Environ. Model. Softw. ***21***(5): 602–614. https://doi.org/10.1016/j.envsoft.2006.01.004.
106. Ma, Z., Chen, M., Yue, S., et al. (2021). Activity-based process construction for participatory geo-analysis. GISci. Remote Sens. ***58***(2): 180–198. https://doi.org/10.1080/15481603.2020.1868211.
107. Gao, J., and O'Neill, B.C. (2020). Mapping global urban land for the 21st century with data-driven simulations and shared socioeconomic pathways. Nat. Commun. ***11***(1): 2302. https://doi.org/10.1038/s41467-020-15788-7.
108. Hey, T. (2009). The Fourth Paradigm (United States of America). https://doi.org/10.1007/s00287-019-01215-9.
109. Liu, X., Chen, M., Claramunt, C., et al. (2022). Geographic information science in the era of geospatial big data: A cyberspace perspective. Innovation ***3***(5): 100279. https://doi.org/10.1016/j.xinn.2022.100279.
110. Wang, S., Han, W., Zhang, X., et al. (2024). Geospatial remote sensing interpretation: From perception to cognition. Innovation Geosci. ***2***(1): 100056–100061. https://doi.org/10.59717/j.xinn-geo.2024.100056.
111. Yang, J., Gong, P., Fu, R., et al. (2013). The role of satellite remote sensing in climate change studies. Nat. Clim. Chang. ***3***(10): 875–883. https://doi.org/10.1038/nclimate1908.
112. Qian, Z., Chen, M., Yang, Y., et al. (2022). Vectorized dataset of roadside noise barriers in China using street view imagery. Earth Syst. Sci. Data ***14***(9): 4057–4076. https://doi.org/10.5194/essd-2022-19.
113. Guo, H., Chen, F., Tang, Y., et al. (2023). Progress toward the sustainable development of world cultural heritage sites facing land-cover changes. Innovation ***4***(5): 100496. https://doi.org/10.1016/j.xinn.2023.100496.

114. Zhang, Z., Qian, Z., Zhong, T., et al. (2022). Vectorized rooftop area data for 90 cities in China. Sci. Data ***9***(1): 66. https://doi.org/10.1038/s41597-022-01168-x.
115. Shi, K., Ma, J., Chen, Z., et al. (2023). Nighttime light remote sensing in characterizing urban spatial structure. Innovation Geosci. ***1***(3): 100043. https://doi.org/10.59717/j.xinn-geo.2023.100043.
116. Gonzalez, M.C., Hidalgo, C.A., and Barabasi, A.-L. (2008). Understanding individual human mobility patterns. Nature ***453***(7196): 779–782. https://doi.org/10.1038/nature07850.
117. Fan, Z., Zhang, F., Loo, B.P.Y., et al. (2023). Urban visual intelligence: Uncovering hidden city profiles with street view images. Proc. Natl. Acad. Sci. USA ***120***(27): e2220417120. https://doi.org/10.1073/pnas.2220417120.
118. Zheng, S., Wang, J., Sun, C., et al. (2019). Air pollution lowers Chinese urbanites' expressed happiness on social media. Nat. Hum. Behav. ***3***(3): 237–243. https://doi.org/10.1038/s41562-018-0521-2.
119. Zhang, Z., Chen, M., Zhong, T., et al. (2023). Carbon mitigation potential afforded by rooftop photovoltaic in China. Nat. Commun. ***14***(1): 2347. https://doi.org/10.1038/s41467-023-38079-3.
120. Wei, K., Ouyang, C., Duan, H., et al. (2020). Reflections on the catastrophic 2020 Yangtze River Basin flooding in southern China. Innovation ***1***(2): 100038. https://doi.org/10.1016/j.xinn.2020.100038.
121. Nohrstedt, D., Hileman, J., Mazzoleni, M., et al. (2022). Exploring disaster impacts on adaptation actions in 549 cities worldwide. Nat. Commun. ***13***(1): 3360. https://doi.org/10.1038/s41467-022-31059-z.
122. Zhang, B., Ouyang, C., Cui, P., et al. (2024). Deep learning for cross-region streamflow and flood forecasting at a global scale. Innovation ***5***(3). https://doi.org/10.1016/j.xinn.2024.100617.
123. Chen, M., Qian, Z., Boers, N., et al. (2023). Iterative integration of deep learning in hybrid earth surface system modelling. Nat. Rev. Earth Environ. ***4***(8): 568–581. https://doi.org/10.1038/s43017-023-00452-7.
124. Luo, Y., Ahlström, A., Allison, S.D., et al. (2016). Toward more realistic projections of soil carbon dynamics by earth system models. Global Biogeochem. Cy. ***30***(1): 40–56. https://doi.org/10.1002/2015gb005239.
125. Lu, D., and Ricciuto, D.M. (2019). Efficient surrogate modeling methods for large-scale earth system models based on machine-learning techniques. Geosci. Model Dev. (GMD) ***12***(5): 1791–1807. https://doi.org/10.5194/gmd-2018-327.
126. Gelbrecht, M., Boers, N., and Kurths, J. (2021). Neural partial differential equations for chaotic systems. New J. Phys. ***23***(4): 043005. https://doi.org/10.5194/egusphere-egu21-8262.
127. Han, J., Jentzen, A., and E, W. (2018). Solving high-dimensional partial differential equations using deep learning. Proc. Natl. Acad. Sci. USA ***115***(34): 8505–8510. https://doi.org/10.1073/pnas.1718942115.
128. Rasp, S., Pritchard, M.S., and Gentine, P. (2018). Deep learning to represent subgrid processes in climate models. Proc. Natl. Acad. Sci. USA ***115***(39): 9684–9689. https://doi.org/10.5194/egusphere-egu2020-13982.
129. Schneider, T., Lan, S., Stuart, A., et al. (2017). Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations. Geophys. Res. Lett. ***44***(24): 12–396. https://doi.org/10.1002/2017gl076101.
130. Qian, Z., Chen, M., Sun, Z., et al. (2024). Simultaneous extraction of spatial and attributional building information across large-scale urban landscapes from high-resolution satellite imagery. Sustain. Cities Soc. ***106***: 105393.
131. Kashinath, K., Mustafa, M., Albert, A., et al. (2021). Physics-informed machine learning: Case studies for weather and climate modelling. Philos. T. R. Soc. A. ***379***(2194): 20200093. https://doi.org/10.1098/rsta.2020.0093.
132. Schütt, K.T., Arbabzadah, F., Chmiela, S., et al. (2017). Quantum-chemical insights from deep tensor neural networks. Nat. Commun. ***8***(1): 13890. https://doi.org/10.1038/ncomms13890.
133. Karpatne, A., Watkins, W., Read, J., et al. (2017). Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling. Preprint at arXiv. https://doi.org/10.1201/9781003143376-15.
134. Guo, H., Liu, Z., Jiang, H., et al. (2017). Big earth data: A new challenge and opportunity for digital earth's development. Int. J. Digit. Earth ***10***(1): 1–12. https://doi.org/10.1080/17538947.2016.1264490.
135. Roh, Y., Heo, G., and Whang, S.E. (2021). A survey on data collection for machine learning: A big data-AI integration perspective. IEEE T. Knowl. Data En. ***33***(4): 1328–1347. https://doi.org/10.1109/tkde.2019.2946162.
136. Tamiminia, H., Salehi, B., Mahdianpari, M., et al. (2020). Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. ISPRS J. Photogramm. Remote Sens. ***164***: 152–170. https://doi.org/10.1016/j.isprsjprs.2020.04.001.
137. Reichstein, M., Camps-Valls, G., Stevens, B., et al. (2019). Deep learning and process understanding for data-driven earth system science. Nature ***566***(7743): 195–204.
138. Xu, L., Chen, N., Chen, Z., et al. (2021). Spatiotemporal forecasting in earth system science: Methods, uncertainties, predictability and future directions. Earth Sci. Rev. ***222***: 103828. https://doi.org/10.1016/j.earscirev.2021.103828.
139. Kumar, P., Debele, S.E., Sahani, J., et al. (2021). Nature-based solutions efficiency evaluation against natural hazards: Modelling methods, advantages and limitations. Sci. Total Environ. ***784***: 147058. https://doi.org/10.1016/j.scitotenv.2021.147058.
140. Gong, Y., Yao, H., and Nallanathan, A. (2024). Intelligent sensing, communication, computation and caching for satellite-ground integrated networks. IEEE Netw. ***38***(4): 9–16. https://doi.org/10.1109/mnet.2024.3413543.
141. Zhou, B., Zhang, S., Xue, R., et al. (2023). A review of space-air-ground integrated remote sensing techniques for atmospheric monitoring. J. For. Environ. ***123***: 3–14. https://doi.org/10.1016/j.jes.2021.12.008.
142. Feng, G., Xia, J., Wang, X., et al. (2024). Assessing the disease burden of air pollution on children and adolescents in China from 1990 to 2019. Innovat. Med. ***2***(1): 100057–100061. https://doi.org/10.59717/j.xinn-med.2024.100057.
143. Awais, M., Li, W., Cheema, M.J.M., et al. (2022). UAV-based remote sensing in plant stress imaging using high-resolution thermal sensor for digital agriculture practices: A meta-review. Int. J. Environ. Sci. Te. 1–18. https://doi.org/10.1007/s13762-021-03801-5.
144. Wang, S., Han, W., Huang, X., et al. (2024). Trustworthy remote sensing interpretation: Concepts, technologies, and applications. ISPRS J. Photogramm. Remote Sens. ***209***: 150–172. https://doi.org/10.1016/j.isprsjprs.2024.02.003.
145. Zhao, Y., Zhu, Z., Chen, B., et al. (2023). Towards parallel intelligence: An interdisciplinary solution for complex systems. Innovation ***4***: 100521. https://doi.org/10.1016/j.xinn.2023.100521.
146. Wang, Z., Zhang, J., Hua, P., et al. (2023). Filling in missing pieces in the co-development of artificial intelligence and environmental science. Innovation Geosci. ***1***(1): 100007–100014. https://doi.org/10.59717/j.xinn-geo.2023.100007.
147. Kadow, C., Hall, D.M., and Ulbrich, U. (2020). Artificial intelligence reconstructs missing climate information. Nat. Geosci. ***13***(6): 408–413. https://doi.org/10.5194/egusphere-egu21-16087.
148. Khanna, S., Liu, P., Zhou, L., et al. (2024). DiffusionSat: A generative foundation model for satellite imagery (ICLR).
149. Tang, D., Cao, X., Hou, X., et al. (2024). CRS-Diff: Controllable generative remote sensing foundation model. Preprint at arXiv. https://doi.org/10.48550/arXiv.2403.11614.
150. DATA, M.R.S.B. (2024). Multimodal artificial intelligence foundation models: Unleashing the power of remote sensing big data in earth observation. Innovation ***2***(1): 100055. https://doi.org/10.59717/j.xinn-geo.2024.100055.
151. Wang, F., Yao, D., Li, Y., et al. (2023). AI-enhanced spatial-temporal data-mining technology: New chance for next-generation urban computing. Innovation ***4***(2): 100405. https://doi.org/10.1016/j.xinn.2023.100405.
152. Hillier, M., Wellmann, F., Brodaric, B., et al. (2021). Three-dimensional structural geological modeling using graph neural networks. Math. Geosci. ***53***(8): 1725–1749. https://doi.org/10.5194/egusphere-egu21-12978.
153. Liu, Z., Feng, R., Wang, L., et al. (2022). Dual learning-based graph neural network for remote sensing image super-resolution. IEEE Trans. Geosci. Remote Sens. ***60***: 1–14. https://doi.org/10.1109/tgrs.2022.3199750.
154. Ayush, K., Uzkent, B., Meng, C., et al. (2021). Geography-aware self-supervised learning. ICCV: 10181–10190. https://doi.org/10.1109/iccv48922.2021.01002.
155. Roy, S.K., Deria, A., Hong, D., et al. (2023). Multimodal fusion transformer for remote sensing image classification. IEEE Trans. Geosci. Remote Sens. ***61***: 1–20. https://doi.org/10.1109/tgrs.2023.3286826.
156. Deng, C., Zhang, T., He, Z., et al. (2024). K2: A foundation language model for geoscience knowledge understanding and utilization. In Proc. 17th ACM Int. Conf. Web Search Data Mining, pp. 161–170. https://doi.org/10.1145/3616855.3635772.
157. Lam, R., Sanchez-Gonzalez, A., Willson, M., et al. (2023). Learning skillful medium-range global weather forecasting. Science ***382***(6677): 1416–1421. https://doi.org/10.1126/science.adi2336.
158. Guo, X., Lao, J., Dang, B., et al. (2024). Skysense: A multi-modal remote sensing foundation model towards universal interpretation for earth observation imagery. CVPR: 27672–27683. https://doi.org/10.1109/lgrs.2009.2034248.
159. Lu, G., Yue, S., Yu, Z., et al. (2023). Ubiquitous geographic information for building digital twins of geographic environments. Innovation Geosci. ***1***(2): 100023. https://doi.org/10.59717/j.xinn-geo.2023.100023.
160. Ma, X. (2023). Data Science for Geoscience: Recent Progress and Future Trends from the Perspective of a Data Life Cycle. https://doi.org/10.31223/x55s4d.
161. Pan, J.Z. (2009). Resource description framework. In Handbook on Ontologies, pp. 71–90. https://doi.org/10.1007/978-3-540-92673-3_3.
162. Wang, K., Yue, P., Yu, D., et al. (2023). OGEScript: An OGC-oriented interoperable script API for online geospatial analysis. In 2023 11th Int. Conf. Agro-Geoinformatics, pp. 1–5. https://doi.org/10.1109/agro-geoinformatics59224.2023.10233317.
163. Zhong, Y., Zheng, Z., Ma, A., et al. (2020). COLOR: Cycling, offline learning, and online representation framework for airport and airplane detection using GF-2 satellite images. IEEE Trans. Geosci. Remote Sens. ***58***(12): 8438–8449. https://doi.org/10.1109/tgrs.2020.2987907.
164. Huang, W., Wang, C., Zhang, R., et al. (2023). Voxposer: Composable 3D value maps for robotic manipulation with language models. Preprint at arXiv. https://doi.org/10.48550/arXiv.2307.05973.
165. Gupta, A., Savarese, S., Ganguli, S., et al. (2021). Embodied intelligence via learning and evolution. Nat. Commun. ***12***(1): 5721. https://doi.org/10.1038/s41467-021-25874-z.
166. Bauer, P., Dueben, P.D., Hoefler, T., et al. (2021). The digital revolution of earth-system science. Nat. Comput. ***1***(2): 104–113. https://doi.org/10.1038/s43588-021-00023-0.
167. Guo, H., Nativi, S., Liang, D., et al. (2020). Big earth data science: An information framework for a sustainable planet. Int. J. Digit. Earth ***13***(7): 743–767. https://doi.org/10.1080/17538947.2020.1743785.
168. Li, X., Feng, M., Ran, Y., et al. (2023). Big data in earth system science and progress towards a digital twin. Nat. Rev. Earth Environ. ***4***(5): 319–332. https://doi.org/10.1038/s43017-023-00409-w.

169. Wang, Q., Li, T., Xu, Y., et al. (2023). How to prevent malicious use of intelligent unmanned swarms? Innovation. ***4***(2). https://doi.org/10.1016/j.xinn.2023.100396.
170. Russell, J.L. (1964). Kepler's laws of planetary motion: 1609-1666. Brit. J. Hist. Sci. ***2***(1): 1–24. https://doi.org/10.1017/s0007087400001813.
171. Wang, Q., Feng, Y., Huang, J., et al. (2023). Large-scale generative simulation artificial intelligence: The next hotspot. Innovation ***4***(6): 100516. https://doi.org/10.1016/j.xinn.2023.100516.
172. Bryant, P., Pozzati, G., and Elofsson, A. (2022). Improved prediction of protein-protein interactions using AlphaFold2. Nat. Commun. ***13***(1): 1265. https://doi.org/10.21203/rs.3.rs-951605/v1.
173. Davies, A., Veličković, P., Buesing, L., et al. (2021). Advancing mathematics by guiding human intuition with AI. Nature ***600***(7887): 70–74. https://doi.org/10.1038/s41586-021-04086-x.
174. Zhang, L., Han, J., Wang, H., et al. (2018). Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics. Phys. Rev. Lett. ***120***(14): 143001. https://doi.org/10.1103/physrevlett.120.143001.
175. Todini, E. (1988). Rainfall-runoff modeling—past, present and future. J. Hydrol. X. ***100***(1-3): 341–352. https://doi.org/10.1016/0022-1694(88)90191-6.
176. Liu, Z., and Todini, E. (2002). Towards a comprehensive physically-based rainfall-runoff model. Hydrol. Earth Syst. Sci. ***6***(5): 859–881. https://doi.org/10.5194/hess-6-859-2002.
177. Shen, C., Laloy, E., Elshorbagy, A., et al. (2018). HESS opinions: Incubating deep-learning-powered hydrologic science advances as a community. Hydrol. Earth Syst. Sci. ***22***(11): 5639–5656. https://doi.org/10.5194/hess-22-5639-2018.
178. Brocca, L., Zhao, W., and Lu, H. (2023). High-resolution observations from space to address new applications in hydrology. Innovation ***4***(3): 100437. https://doi.org/10.1016/j.xinn.2023.100437.
179. Wang, H., Fu, T., Du, Y., et al. (2023). Scientific discovery in the age of artificial intelligence. Nature ***620***(7972): 47–60.
180. Marino, K., Fergus, R., Szlam, A., et al. (2020). Empirically verifying hypotheses using reinforcement learning. Preprint at arXiv. https://doi.org/10.48550/arXiv.2006.15762.
181. Petersen, B.K., Landajuela, M., Mundhenk, T.N., et al. (2019). Deep symbolic regression: Recovering mathematical expressions from data via risk-seeking policy gradients. Preprint at arXiv. https://doi.org/10.48550/arXiv.1912.04871.
182. Flecker, A.S., Shi, Q., Almeida, R.M., et al. (2022). Reducing adverse impacts of Amazon hydropower expansion. Science ***375***(6582): 753–760. https://doi.org/10.1126/science.abj4017.
183. Xue, Y., Palmer-Brown, D., and Guo, H. (2011). The use of high-performance and high-throughput computing for the fertilization of digital earth and global change studies. Int. J. Digit. Earth ***4***(3): 185–210. https://doi.org/10.1080/17538947.2010.535569.
184. Brunton, S.L., Proctor, J.L., and Kutz, J.N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proc. Natl. Acad. Sci. USA ***113***(15): 3932–3937. https://doi.org/10.1073/pnas.1517384113.
185. Stupp, D., Sharon, E., Bloch, I., et al. (2021). Co-evolution based machine-learning for predicting functional interactions between human genes. Nat. Commun. ***12***(1): 6454. https://doi.org/10.1038/s41467-021-26792-w.
186. Kalra, S., Lamba, R., and Sharma, M. (2020). Machine learning based analysis for relation between global temperature and concentrations of greenhouse gases. J. Inf. Optim. Sci. ***41***(1): 73–84. https://doi.org/10.1080/02522667.2020.1715559.
187. Bi, K., Xie, L., Zhang, H., et al. (2023). Accurate medium-range global weather forecasting with 3D neural networks. Nature ***619***(7970): 533–538. https://doi.org/10.1038/s41586-023-06185-3.
188. Beer, C., Reichstein, M., Tomelleri, E., et al. (2010). Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. Science ***329***(5993): 834–838. https://doi.org/10.1126/science.1184984.
189. Forkel, M., Carvalhais, N., Rödenbeck, C., et al. (2016). Enhanced seasonal CO2 exchange caused by amplified plant productivity in northern ecosystems. Science ***351***(6274): 696–699. https://doi.org/10.1126/science.aac4971.
190. Wu, L., Wang, L., Li, N., et al. (2020). Modeling the COVID-19 outbreak in China through multi-source information fusion. Innovation ***1***(2): 100033. https://doi.org/10.1016/j.xinn.2020.100033.
191. Burns, J.A. (2010). The four hundred years of planetary science since Galileo and Kepler. Nature ***466***(7306): 575–584. https://doi.org/10.1038/nature09215.
192. Greene, M.T. (2015). Alfred Wegener: Science, Exploration, and the Theory of Continental Drift (JHU Press). https://doi.org/10.5860/choice.195372.
193. Zou, Z., Xiao, X., Dong, J., et al. (2018). Divergent trends of open-surface water body area in the contiguous United States from 1984 to 2016. Proc. Natl. Acad. Sci. USA ***115***(15): 3810–3815. https://doi.org/10.1073/pnas.1719275115.
194. Bastin, J.F., Finegold, Y., Garcia, C., et al. (2019). The global tree restoration potential. Science ***365***(6448): 76–79. https://doi.org/10.1126/science.aaz0493.
195. Yu, C., Yan, G., Yu, C., et al. (2023). A multi-factor driven spatiotemporal wind power prediction model based on ensemble deep graph attention reinforcement learning networks. Energy ***263***: 126034. https://doi.org/10.1016/j.energy.2022.126034.
196. Yu, C., Yan, G., Yu, C., et al. (2023). Attention mechanism is useful in spatio-temporal wind speed prediction: Evidence from China. Appl. Soft Comput. ***148***: 110864. https://doi.org/10.1016/j.asoc.2023.110864.
197. Horie, M., and Mitsume, N. (2022). Physics-embedded neural networks: Graph neural PDE solvers with mixed boundary conditions. Adv. Neural Inf. Process. Syst. ***35***: 23218–23229. https://doi.org/10.1299/jsmecmd.2022.35.16-16.
198. Ma, M., Xie, P., Teng, F., et al. (2023). HistGNN: Hierarchical spatiotemporal graph neural network for weather forecasting. Inf. Sci. ***648***: 119580. https://doi.org/10.2139/ssrn.4455568.
199. Shao, Z., Zhang, Z., Wang, F., et al. (2022). Pre-training enhanced spatial-temporal graph neural network for multivariate time series forecasting. In Proc. 28th ACM SIGKDD Conf. Knowl. Discov. Data Min, pp. 1567–1577. https://doi.org/10.1145/3534678.3539396.
200. Yu, C., Wang, F., Shao, Z., et al. (2023). DSFormer: A double sampling transformer for multivariate time series long-term prediction. In Proc. 32nd ACM Int. Conf. Inf. Knowl. Manag, pp. 3062–3072. https://doi.org/10.1145/3583780.3614851.
201. Giladi, N., Ben-Haim, Z., Nevo, S., et al. (2021). Physics-aware downsampling with deep learning for scalable flood modeling. Adv. Neural Inf. Process. Syst. ***34***: 1378–1389. https://doi.org/10.5194/hess-2021-554.
202. Walter, T.R., Haghshenas Haghighi, M., Schneider, F.M., et al. (2019). Complex hazard cascade culminating in the Anak Krakatau sector collapse. Nat. Commun. ***10***(1): 4339. https://doi.org/10.1038/s41467-019-12284-5.
203. Yu, H., Zahidi, I., Fai, C.M., et al. (2024). Breathing in the new era: The global call against industrial air pollution. Innovation Med. ***2***(1): 100049. https://doi.org/10.59717/j.xinn-med.2024.100049.
204. Klemmer, K., Rolf, E., Robinson, C., et al. (2023). Satclip: Global, general-purpose location embeddings with satellite imagery. Preprint at arXiv. https://doi.org/10.1038/d41586-023-03983-7.
205. Chen, F., Xue, G., Wang, Y., et al. (2023). Evolution of the Yangtze River and its biodiversity. Innovation ***4***(3): 100417.
206. Zhang, P., Goodman, S.J., Bai, S., et al. (2023). Marine mammal genomes: Important resources for unraveling adaptation and evolution in the marine environment. Innovation Geosci. ***1***(2). https://doi.org/10.59717/j.xinn-geo.2023.100022.
207. Qi, J., Wang, Y., Wang, L., et al. (2023). The modification effect of ozone pollution on the associations between heat wave and cardiovascular mortality. Innovation ***1***(3): 100043. https://doi.org/10.59717/j.xinn-med.2023.100043.
208. Sun, Y., Wang, M., Wang, Y., et al. (2023). Exposure to airborne PM2.5 chemical exposome increases heart rate of middle- and old-aged populations. Innovation Med. ***1***(10.59717). https://doi.org/10.59717/j.xinn-med.2023.100042.
209. Du, J., and Feng, X. (2024). Biomedical microrobotics: Small sizes, large applications. Innovation Life ***2***(1): 100046. https://doi.org/10.59717/j.xinn-life.2024.100046.
210. Ahmed, A.A.M., Jui, S.J.J., Sharma, E., et al. (2024). An advanced deep learning predictive model for air quality index forecasting with remote satellite-derived hydro-climatological variables. Sci. Total Environ. ***906***: 167234. https://doi.org/10.1016/j.scitotenv.2023.167234.
211. Li, S., and Xing, J. (2024). DeepSat4D: Deep learning empowers four-dimensional atmospheric chemical concentration and emission retrieval from satellite. Innovation Geosci. ***2***(1): 100061. https://doi.org/10.59717/j.xinn-geo.2024.100061.
212. Wu, Y., Gasevic, D., Wen, B., et al. (2023). Association between air pollution and telomere length: A study of 471,808 UK Biobank participants. Innovat. Med. ***1***(2): 100017. https://doi.org/10.59717/j.xinn-med.2023.100017.
213. Le Goff, C., Fablet, R., Tandeo, P., et al. (2016). Spatio-temporal decomposition of satellite-derived SST-SSH fields: Links between surface data and ocean interior dynamics in the Agulhas region. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. ***9***(11): 5106–5112. https://doi.org/10.1109/jstars.2016.2605040.
214. Chen, Y., Haywood, J., Wang, Y., et al. (2022). Machine learning reveals climate forcing from aerosols is dominated by increased cloud cover. Nat. Geosci. ***15***(8): 609–614. https://doi.org/10.1038/s41561-022-01027-9.
215. Yang, Y., Sun, W., Chi, Y., et al. (2022). Machine learning-based retrieval of day and night cloud macrophysical parameters over East Asia using Himawari-8 data. Remote Sens. Environ. ***273***: 112971. https://doi.org/10.1016/j.rse.2022.112971.
216. Wang, L., Dong, H., Cao, Y., et al. (2023). Real-time water quality detection based on fluctuation feature analysis with the LSTM model. J. Hydroinform. ***25***(1): 140–149. https://doi.org/10.2166/hydro.2023.127.
217. Mahajan, S., and Fataniya, B. (2020). Cloud detection methodologies: Variants and development—a review. Complex Intell. Syst. ***6***: 251–261. https://doi.org/10.1007/s40747-019-00128-0.
218. Jiang, N., Xu, Y., Xu, T., et al. (2022). Land water vapor retrieval for AMSR2 using a deep learning method. IEEE Trans. Geosci. Remote Sens. ***60***: 1–11. https://doi.org/10.5194/egusphere-egu2020-17756.
219. Di Noia, A., and Hasekamp, O.P. (2018). Neural Networks and Support Vector Machines and Their Application to Aerosol and Cloud Remote Sensing: A Review (Springer Series in Light Scattering), pp. 279–329. https://doi.org/10.1007/978-3-319-70796-9_4.
220. Kang, Y., Kim, M., Kang, E., et al. (2022). Improved retrievals of aerosol optical depth and fine mode fraction from GOCI geostationary satellite data using machine learning over East Asia. ISPRS J. Photogramm. Remote Sens. ***183***: 253–268. https://doi.org/10.1016/j.isprsjprs.2021.11.016.
221. Shi, C., Hashimoto, M., Shiomi, K., et al. (2020). Development of an algorithm to retrieve aerosol optical properties over water using an artificial neural network radiative transfer scheme: First result from GOSAT-2/CAI-2. IEEE Trans. Geosci. Remote Sens. ***59***(12): 9861–9872. https://doi.org/10.1109/tgrs.2020.3038892.
222. Su, T., Laszlo, I., Li, Z., et al. (2020). Refining aerosol optical depth retrievals over land by constructing the relationship of spectral surface reflectances through deep learning: Application to Himawari-8. Remote Sens. Environ. ***251***: 112093. https://doi.org/10.1016/j.rse.2020.112093.

223. Liu, T., Meng, H., Yu, M., et al. (2021). Urban-rural disparity of the short-term association of PM2.5 with mortality and its attributable burden. Innovation **2**(4): 100171. https://doi.org/10.1016/j.xinn.2021.100171.
224. Liu, S., Geng, G., Xiao, Q., et al. (2022). Tracking daily concentrations of PM2.5 chemical composition in China since 2000. Environ. Sci. Technol. **56**(22): 16517–16527. https://doi.org/10.1021/acs.est.2c06510.
225. Wei, J., Li, Z., Chen, X., et al. (2023). Separating daily 1 km PM2.5 inorganic chemical composition in China since 2000 via deep learning integrating ground, satellite, and model data. Environ. Sci. Technol. **57**(46): 18282–18295. https://doi.org/10.1021/acs.est.3c00272.
226. Nguyen, T.-A., Kellenberger, B., and Tuia, D. (2022). Mapping forest in the Swiss Alps treeline ecotone with explainable deep learning. Remote Sens. Environ. **281**: 113217. https://doi.org/10.1016/j.rse.2022.113217.
227. Jerrett, M., Arain, A., Kanaroglou, P., et al. (2005). A review and evaluation of intraurban air pollution exposure models. J. Expo. Sci. Environ. Epidemiol. **15**(2): 185–204. https://doi.org/10.1038/sj.jea.7500388.
228. Kang, Y., Choi, H., Im, J., et al. (2021). Estimation of surface-level NO2 and O3 concentrations using TROPOMI data and machine learning over East Asia. Environ. Pollut. **288**: 117711. https://doi.org/10.1016/j.envpol.2021.117711.
229. Wang, F., Xiang, L., Leung, K.S.-Y., et al. (2024). Emerging contaminants: A One Health perspective. Innovation **5**(4): 2666–2758. https://doi.org/10.1016/j.xinn.2024.100612.
230. Sun, K., Jia, J., Wang, S., et al. (2023). Real-time and dynamic estimation of CO2 emissions from China's lakes and reservoirs. Innovation Geosci. **1**(3): 100031. https://doi.org/10.59717/j.xinn-geo.2023.100031.
231. Koç, E., Zeybek, S., Özbay Kısasöz, B., et al. (2022). Estimation of surface roughness in selective laser sintering using computational models. Int. J. Adv. Manuf. Technol. **123**(9-10): 3033–3045. https://doi.org/10.21203/rs.3.rs-1638732/v1.
232. Letu, H., Ma, R., Nakajima, T.Y., et al. (2023). Surface solar radiation compositions observed from Himawari-8/9 and Fengyun-4 series. B Am Meteorol Soc **104**(10): E1772–E1789. https://doi.org/10.1175/bams-d-22-0154.1.
233. Zhang, Y., Long, M., Chen, K., et al. (2023). Skilful nowcasting of extreme precipitation with NowcastNet. Nature **619**(7970): 526–532. https://doi.org/10.1038/s41586-023-06184-4.
234. Hu, Y., Chen, L., Wang, Z., et al. (2023). SwinVRNN: A data-driven ensemble forecasting model via learned distribution perturbation. J. Adv. Model. Earth Syst. **15**(2): e2022MS003211. https://doi.org/10.1029/2022ms003211.
235. Chen, L., Zhong, X., Zhang, F., et al. (2023). FuXi: A cascade machine learning forecasting system for 15-day global weather forecast. NPJ Clim. Atmos. Sci. **6**(1): 190. https://doi.org/10.1038/s41612-023-00512-1.
236. Chen, K., Han, T., Gong, J., et al. (2023). Fengwu: Pushing the skillful global medium-range weather forecast beyond 10 days lead. Preprint at arXiv. https://doi.org/10.48550/arXiv.2304.02948.
237. Singh, M., Acharya, N., Patel, P., et al. (2023). A modified deep learning weather prediction using cubed sphere for global precipitation. Front. Clim. **4**: 1022624. https://doi.org/10.3389/fclim.2022.1022624.
238. Verendel, V. (2023). Tracking artificial intelligence in climate inventions with patent data. Nat. Clim. Chang. **13**(1): 40–47. https://doi.org/10.1038/s41558-022-01536-w.
239. Christensen, N.I., and Mooney, W.D. (1995). Seismic velocity structure and composition of the continental crust: A global view. J. Geophys. Res. **100**(B6): 9761–9788. https://doi.org/10.1029/94JB03148.
240. Shi, D. (2023). Morning twilight of crop breeding for sodic land. Innov. Life **1**(2): 100020–100021.
241. Tariq, Z., Aljawad, M.S., Hasan, A., et al. (2021). A systematic review of data science and machine learning applications to the oil and gas industry. J. Pet. Explor. Prod. Technol. **1**: 1–36. https://doi.org/10.1007/s13202-021-01302-2.
242. Huang, R., Liu, S., Qi, R., et al. (2021). Deep learning 3D sparse inversion of gravity data. JGR. Solid Earth **126**(11): e2021JB022476. https://doi.org/10.1029/2021jb022476.
243. Chen, Y., and Saygin, E. (2021). Seismic inversion by hybrid machine learning. JGR. Solid Earth **126**(9): e2020JB021589. https://doi.org/10.1029/2020jb021589.
244. Xie, L., Han, B., Hu, X., et al. (2023). 2D magnetotelluric inversion based on ResNet. Artif. Intell. Geosci **4**: 119–127. https://doi.org/10.1016/j.aiig.2023.08.003.
245. Zhang, L., Zhang, G., Liu, Y., et al. (2021). Deep learning for 3-D inversion of gravity data. IEEE Trans. Geosci. Remote Sens. **60**: 1–18. https://doi.org/10.1109/tgrs.2021.3110606.
246. Cai, D., Aziz, G., Sarwar, S., et al. (2024). Applicability of denoising-based artificial intelligence to forecast the environmental externalities. Geosci. Front. **15**(3): 101740. https://doi.org/10.1016/j.gsf.2023.101740.
247. Zhu, L., Peng, Z., McClellan, J., et al. (2019). Deep learning for seismic phase detection and picking in the aftershock zone of 2008 Mw 7.9 Wenchuan earthquake. Phys. Earth Planet. Inter. **293**: 106261. https://doi.org/10.1016/j.pepi.2020.106261.
248. Meng, F., Ren, T., Liu, Z., et al. (2023). Toward earthquake early warning: A convolutional neural network for rapid earthquake magnitude estimation. Artif. Intell. Geosci. **4**: 39–46. https://doi.org/10.1016/j.aiig.2023.03.001.
249. Kong, Q., Trugman, D.T., Ross, Z.E., et al. (2019). Machine learning in seismology: Turning data into insights. Seismol Res. Lett. **90**(1): 3–14. https://doi.org/10.1785/0220180259.
250. Md Ridzwan, N.S., and Md Yusoff, S.H. (2023). Machine learning for earthquake prediction: A review (2017-2021). Earth Sci. Inform. **16**(2): 1133–1149. https://doi.org/10.1007/s12145-023-00991-z.
251. Youssef, K., Shao, K., Moon, S., et al. (2023). Landslide susceptibility modeling by interpretable neural network. Commun. Earth Environ. **4**(1): 162. https://doi.org/10.1038/s43247-023-00806-5.
252. Anantrasirichai, N., Biggs, J., Albino, F., et al. (2018). Application of machine learning to classification of volcanic deformation in routinely generated InSAR data. JGR. Solid Earth **123**(8): 6592–6606. https://doi.org/10.1029/2018jb015911.
253. Laurenti, L., Tinti, E., Galasso, F., et al. (2022). Deep learning for laboratory earthquake prediction and autoregressive forecasting of fault zone stress. Earth Planet Sci. Lett. **598**: 117825. https://doi.org/10.1016/j.epsl.2022.117825.
254. Pwavodi, J., Ibrahim, A.U., Pwavodi, P.C., et al. (2024). The role of artificial intelligence and IoT in prediction of earthquakes. Artif. Intell. Geosci. **4**: 100075. https://doi.org/10.1016/j.aiig.2024.100075.
255. Lawal, A.I., and Kwon, S. (2021). Application of artificial intelligence to rock mechanics: An overview. J. Rock Mech. Geotech. **13**(1): 248–266. https://doi.org/10.1016/j.jrmge.2020.05.010.
256. Alzubaidi, F., Mostaghimi, P., Swietojanski, P., et al. (2021). Automated lithology classification from drill core images using convolutional neural networks. J. Pet. Sci. Eng. **197**: 107933. https://doi.org/10.1016/j.petrol.2020.107933.
257. He, S., Cai, H., Liu, S., et al. (2021). Recovering 3D basement relief using gravity data through convolutional neural networks. JGR. Solid Earth **126**(10): e2021JB022611. https://doi.org/10.1029/2021jb022611.
258. Cui, Z., Chen, Q., and Liu, G. (2022). Characterization of subsurface hydrogeological structures with convolutional conditional neural processes on limited training data. Water Resour. Res. **58**(12): e2022WR033161. https://doi.org/10.1029/2022wr033161.
259. Wang, S., Cai, Z., Si, X., et al. (2023). A three-dimensional geological structure modeling framework and its application in machine learning. Math. Geosci. **55**(2): 163–200. https://doi.org/10.1007/s11004-022-10027-9.
260. Padarian, J., Minasny, B., and McBratney, A.B. (2020). Machine learning and soil sciences: A review aided by machine learning tools. Soil **6**(1): 35–52. https://doi.org/10.5194/soil-6-35-2020.
261. Orth, R. (2021). Global soil moisture data derived through machine learning trained with in-situ measurements. Sci. Data **8**(1): 1170. https://doi.org/10.1038/s41597-021-00964-1.
262. Vakilzadeh Ebrahimi, M.K., Lee, H., Won, J., et al. (2023). Estimation of soil texture by fusion of near-infrared spectroscopy and image data based on convolutional neural network. Comput. Electron. Agric. **212**: 108117. https://doi.org/10.1016/j.compag.2023.108117.
263. de Andrade, V.H.G.Z., Redmile-Gordon, M., Barbosa, B.H.G., et al. (2021). Artificially intelligent soil quality and health indices for 'next generation' food production systems. Trends Food Sci. Technol. **107**: 195–200. https://doi.org/10.1016/j.tifs.2020.10.018.
264. Bailo, D., Paciello, R., Michalek, J., et al. (2023). The EPOS multi-disciplinary data portal for integrated access to solid earth science datasets. Sci. Data **10**(1): 784. https://doi.org/10.1038/s41597-023-02697-9.
265. Araya, Y.N. (2015). Hydrosphere. https://doi.org/10.1002/047147844x.me216.
266. Mosaffa, H., Sadeghi, M., Mallakpour, I., et al. (2022). Application of machine learning algorithms in hydrology. Comput. Earth Environ **1**: 585–591. https://doi.org/10.1016/b978-0-323-89861-4.00027-0.
267. Liu, K., Gao, W., Xu, W., et al. (2023). Deep-sea microorganisms acquired during Jiaolong expedition. Innov. Life **1**(2): 100029.
268. Yang, N., and Li, X. (2024). Lightweight AI-powered precipitation nowcasting. The Innovation Geoscience **2**(2): 100066. https://doi.org/10.59717/j.xinn-geo.2024.100066.
269. Luo, C., Li, X., Ye, Y., et al. (2022). Experimental study on generative adversarial network for precipitation nowcasting. IEEE Trans. Geosci. Remote Sens. **60**: 1–20. https://doi.org/10.59717/j.xinn-geo.2024.100066.
270. Chen, H., Chandrasekar, V., Cifelli, R., et al. (2019). A machine learning system for precipitation estimation using satellite and ground radar network observations. IEEE Trans. Geosci. Remote Sens. **58**(2): 982–994. https://doi.org/10.1109/tgrs.2019.2942280.
271. Espeholt, L., Agrawal, S., Sønderby, C., et al. (2022). Deep learning for twelve hour precipitation forecasts. Nat. Commun. **13**(1): 1–10. https://doi.org/10.1038/s41467-022-32483-x.
272. Zhang, X., Song, Y., Nam, W.-H., et al. (2024). Data fusion of satellite imagery and downscaling for generating highly fine-scale precipitation. J. Hydrol. X. **631**: 130665. https://doi.org/10.1016/j.jhydrol.2024.130665.
273. Davenport, F.V., and Diffenbaugh, N.S. (2021). Using machine learning to analyze physical causes of climate change: A case study of US Midwest extreme precipitation. Geophys. Res. Lett. **48**(15): e2021GL093787. https://doi.org/10.1029/2021GL093787.
274. Dong, J., Zhu, Y., Jia, X., et al. (2022). Nation-scale reference evapotranspiration estimation by using deep learning and classical machine learning models in China. J. Hydrol. X. **604**: 127207.
275. Pan, S., Pan, N., Tian, H., et al. (2020). Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning, and land surface modeling. Hydrol. Earth Syst. Sci. **24**(3): 1485–1509. https://doi.org/10.5194/hess-2019-409-supplement.
276. Xu, T., Guo, Z., Liu, S., et al. (2018). Evaluating different machine learning methods for upscaling evapotranspiration from flux towers to the regional scale. JGR. Atmospheres **123**(16): 8674–8690. https://doi.org/10.1029/2018JD028447.
277. Torres, A.F., Walker, W.R., and McKee, M. (2011). Forecasting daily potential evapotranspiration using machine learning and limited climatic data. Agric. Water Manag. **98**(4): 553–562. https://doi.org/10.1016/j.agwat.2010.10.012.
278. Granata, F., and Di Nunno, F. (2021). Forecasting evapotranspiration in different climates using ensembles of recurrent neural networks. Agric. Water Manag. **255**: 107040. https://doi.org/10.1016/j.agwat.2021.107040.

279. Kumar, A., Ramsankaran, R.A.A.J., Brocca, L., et al. (2021). A simple machine learning approach to model real-time streamflow using satellite inputs: Demonstration in a data scarce catchment. J. Hydrol. X. ***595***: 126046. https://doi.org/10.1016/j.jhydrol.2021.126046.
280. Rasouli, K., Hsieh, W.W., and Cannon, A.J. (2012). Daily streamflow forecasting by machine learning methods with weather and climate inputs. J. Hydrol. X. ***414–415***: 284–293. https://doi.org/10.1016/j.jhydrol.2011.10.039.
281. Ferreira, R.G., Silva, D.D., Alden Elesbon, A.A., et al. (2021). Machine learning models for streamflow regionalization in a tropical watershed. J. Environ. ***280***: 111713. https://doi.org/10.1016/j.jenvman.2020.111713.
282. Hagen, J.S., Leblois, E., Lawrence, D., et al. (2021). Identifying major drivers of daily streamflow from large-scale atmospheric circulation with machine learning. J. Hydrol. X. ***596***: 126086. https://doi.org/10.1016/j.jhydrol.2021.126086.
283. Konapala, G., Kao, S.-C., Painter, S.L., et al. (2020). Machine learning assisted hybrid models can improve streamflow simulation in diverse catchments across the conterminous US. Environ. Res. Lett. ***15***(10): 104022. https://doi.org/10.1088/1748-9326/aba927.
284. Chen, Y., Tang, L., Kan, Z., et al. (2020). A novel water body extraction neural network (WBE-NN) for optical high-resolution multispectral imagery. J. Hydrol. ***588***: 125092. https://doi.org/10.1016/j.jhydrol.2020.125092.
285. Nyberg, B., Henstra, G., Gawthorpe, R.L., et al. (2023). Global scale analysis on the extent of river channel belts. Nat. Commun. ***14***(1): 2163. https://doi.org/10.1038/s41467-023-37852-8.
286. Tiyasha, Tung, T.M., and Yaseen, Z.M. (2020). A survey on river water quality modelling using artificial intelligence models: 2000-2020. J. Hydrol. ***585***: 124670. https://doi.org/10.1016/j.jhydrol.2020.124670.
287. Do, H.X., Gudmundsson, L., Leonard, M., et al. (2018). The global streamflow indices and meta-data archive (GSIM)-Part 1: The production of a daily streamflow archive and metadata. Earth Syst. Sci. Data ***10***(2): 765–785. https://doi.org/10.5194/essd-10-765-2018.
288. Harrigan, S., Zsoter, E., Alfieri, L., et al. (2020). GLOFAS-ERA5 operational global river discharge reanalysis 1979-present. Earth Syst. Sci. Data ***12***(3): 2043–2060. https://doi.org/10.5194/essd-12-2043-2020.
289. Hartmann, J., Lauerwald, R., and Moosdorf, N. (2014). A brief overview of the global river chemistry database, GLORICH. Prog. Earth Planet. Sci. ***10***: 23–27. https://doi.org/10.1016/j.proeps.2014.08.005.
290. Virro, H., Amatulli, G., Kmoch, A., et al. (2021). GRQA: Global river water quality archive. Earth Syst. Sci. Data ***13***(12): 5483–5507. https://doi.org/10.5194/egusphere-egu21-3865.
291. Cheng, Y., and Liu, H. (2024). The crucial role of soil moisture in the evolution of forest cover in Asia since the last glacial maximum. Innovation. ***5***(3): 100594. https://doi.org/10.1016/j.xinn.2024.100594.
292. Rodriguez-Fernandez, N., Aires, F., Richaume, P., et al. (2015). Soil moisture retrieval using neural networks: Application to SMOS. IEEE Trans. Geosci. Remote Sens. ***53***: 5991–6007. https://doi.org/10.1109/tgrs.2015.2430845.
293. Kolassa, J., Gentine, P., Prigent, C., et al. (2017). Soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 2: Product evaluation. Remote Sens. Environ. ***195***: 202–217. https://doi.org/10.1016/j.rse.2017.04.020.
294. Ge, L., Hang, R., and Liu, Q. (2019). Retrieving soil moisture over continental US via multi-view multi-task learning. IEEE Geosci. Remote Sens. Lett. ***16***(12): 1954–1958. https://doi.org/10.1109/lgrs.2019.2913100.
295. Abowarda, A.S., Bai, L., Zhang, C., et al. (2021). Generating surface soil moisture at 30 m spatial resolution using both data fusion and machine learning toward better water resources management at the field scale. Remote Sens. Environ. ***255***: 112301. https://doi.org/10.1016/j.rse.2021.112301.
296. Liu, Y., Jing, W., Wang, Q., et al. (2020). Generating high-resolution daily soil moisture by using spatial downscaling techniques: A comparison of six machine learning algorithms. Adv. Water Resour. ***141***: 103601. https://doi.org/10.1016/j.advwatres.2020.103601.
297. Xu, W., Zhang, Z., Long, Z., et al. (2021). Downscaling SMAP soil moisture products with convolutional neural network. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. ***14***: 4051–4062. https://doi.org/10.1109/jstars.2021.3069774.
298. Zheng, C., Jia, L., and Zhao, T. (2023). A 21-year dataset (2000-2020) of gap-free global daily surface soil moisture at 1-km grid resolution. Sci. Data ***10***(1): 139. https://doi.org/10.1038/s41597-023-01991-w.
299. Huang, S., Zhang, X., Chen, N., et al. (2022). Generating high-accuracy and cloud-free surface soil moisture at 1 km resolution by point-surface data fusion over the southwestern US. Agric. For. Meteorol. ***321***: 108985. https://doi.org/10.1016/j.agrformet.2022.108985.
300. Yao, P., Lu, H., Shi, J., et al. (2021). A long-term global daily soil moisture dataset derived from AMSR-E and AMSR2 (2002-2019). Sci. Data ***8***(1): 143. https://doi.org/10.1038/s41597-021-00925-8.
301. Yao, P., Lu, H., Zhao, T., et al. (2023). A global daily soil moisture dataset derived from Chinese FengYun Microwave Radiation Imager (MWRI) (2010-2019). Sci. Data. ***10***(1): 133. https://doi.org/10.5194/egusphere-egu2020-311.
302. Babaeian, E., Paheding, S., Siddique, N., et al. (2021). Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning. Remote Sens. Environ. ***260***: 112434. https://doi.org/10.1016/j.rse.2021.112434.
303. Zhu, Z., Zhao, C., Jia, X., et al. (2023). Prediction of deep soil water content (0-5 m) with in-situ and remote sensing data. Catena ***222***: 106852. https://doi.org/10.1016/j.catena.2022.106852.
304. Li, P., Zha, Y., Shi, L., et al. (2020). Comparison of the use of a physical-based model with data assimilation and machine learning methods for simulating soil water dynamics. J. Hydrol. X. ***584***: 124692. https://doi.org/10.1016/j.jhydrol.2020.124692.
305. Zhang, Q., Shi, L., Holzman, M., et al. (2019). A dynamic data-driven method for dealing with model structural error in soil moisture data assimilation. Adv. Water Resour. ***132***: 103407. https://doi.org/10.1016/j.advwatres.2019.103407.
306. Sun, A.Y., Scanlon, B.R., Save, H., et al. (2021). Reconstruction of GRACE total water storage through automated machine learning. Water Resour. Res. ***57***(2): e2020WR028666. https://doi.org/10.5194/gstm2020-53.
307. Yin, J., Slater, L.J., Khouakhi, A., et al. (2023). LREC: Global terrestrial water storage reconstruction by machine learning from 1940 to present. Earth Syst. Sci. Data ***2023***: 1–29. https://doi.org/10.5194/essd-2023-315-supplement.
308. Nketia, K.A., Asabere, S.B., Ramcharan, A., et al. (2022). Temporal mapping of soil water storage in a semi-arid landscape of northern Ghana - a multi-tasked ensemble machine-learning approach. Geoderma ***410***: 115691. https://doi.org/10.1016/j.geoderma.2021.115691.
309. Foroumandi, E., Nourani, V., Huang, J.J., et al. (2023). Monitoring by downscaling GRACE-derived terrestrial water storage anomalies: A deep learning approach. J. Hydrol. ***616***: 128838. https://doi.org/10.1016/j.jhydrol.2022.128838.
310. He, H., Yang, K., Wang, S., et al. (2021). Learning approaches to spatial downscaling of GRACE terrestrial water storage products using EALCO model over Canada. Can. J. Remote Sens. ***47***(4): 657–675. https://doi.org/10.1080/07038992.2021.1954498.
311. Kagemoto, H. (2022). Forecasting a water-surface wave train with artificial intelligence (Part 2)-Can the occurrence of freak waves be predicted with AI? Ocean Eng ***252***: 111205. https://doi.org/10.1016/j.oceaneng.2022.111205.
312. Katija, K., Orenstein, E., Schlining, B., et al. (2022). FathomNet: A global image database for enabling artificial intelligence in the ocean. Sci. Rep. ***12***(1): 15914. https://doi.org/10.1038/s41598-022-19939-2.
313. Hübscher, C., and Nürnberg, D. (2023). Loop current attenuation after the mid-Pleistocene transition contributes to northern hemisphere cooling. Mar. Geol. ***456***: 106976. https://doi.org/10.1016/j.margeo.2022.106976.
314. Zhang, H., Huang, B., Chen, G., et al. (2022). An efficient oceanic eddy identification method with XBT data using transformer. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. ***15***: 9860–9872. https://doi.org/10.1109/jstars.2022.3221113.
315. Zhang, L., Zhang, Y., and Yin, X. (2023). Aquarius sea surface salinity retrieval in coastal regions based on deep neural networks. Remote Sens. Environ. ***284***: 113357. https://doi.org/10.1016/j.rse.2022.113357.
316. Jang, E., Kim, Y.J., Im, J., et al. (2021). Improvement of SMAP sea surface salinity in river-dominated oceans using machine learning approaches. GISci. Remote Sens. ***58***(1): 138–160. https://doi.org/10.1080/15481603.2021.1872228.
317. Kim, D.-W., Kim, S.-H., Baek, J.-Y., et al. (2022). GOCI-II based sea surface salinity estimation using machine learning for the first-year summer. Int. J. Remote Sens. ***43***(18): 6605–6623. https://doi.org/10.1080/01431161.2022.2142080.
318. Meng, L., Yan, C., Zhuang, W., et al. (2021). Reconstruction of three-dimensional temperature and salinity fields from satellite observations. JGR. Oceans ***126***(11): e2021JC017605. https://doi.org/10.1029/2021jc017605.
319. Jang, E., Kim, Y.J., Im, J., et al. (2022). Global sea surface salinity via the synergistic use of SMAP satellite and HYCOM data based on machine learning. Remote Sens. Environ. ***273***: 112980. https://doi.org/10.1016/j.rse.2022.112980.
320. Peng, C., Zeng, J., Chen, K.-S., et al. (2023). Global spatiotemporal trend of satellite-based soil moisture and its influencing factors in the early 21st century. Remote Sens. Environ. ***291***: 113569. https://doi.org/10.1016/j.rse.2023.113569.
321. Dikshit, A., Pradhan, B., Matin, S.S., et al. (2024). Artificial intelligence: A new era for spatial modelling and interpreting climate-induced hazard assessment. Geosci. Front. ***15***: 101815.
322. Xu, Q., Shi, Y., Bamber, J., et al. (2024). Large-scale flood modeling and forecasting with Floodcast. Water Res.. Preprint at arXiv. https://doi.org/10.1016/j.watres.2024.122162.
323. Başagaoğlu, H., Chakraborty, D., Do Lago, C., et al. (2022). A review on interpretable and explainable artificial intelligence in hydroclimatic applications. Water ***14***(8): 1230. https://doi.org/10.3390/w14081230.
324. Xu, Q., Shi, Y., Bamber, J., et al. (2023). Physics-aware machine learning revolutionizes scientific paradigm for machine learning and process-based hydrology. Preprint at arXiv. https://doi.org/10.48550/arXiv.2310.05227.
325. Ndehedehe, C. (2023). Hydro-Climatic Extremes in the Anthropocene (Springer Int. Publ. AG).
326. Zheng, G., Muhammad, S., Sattar, A., et al. (2023). Cryosphere remote sensing. Front. Remote Sens. ***4***: 1204667. https://doi.org/10.3389/frsen.2023.1204667.
327. Slaymaker, O., and Kelly, R. (2009). The cryosphere and global environmental change (John Wiley & Sons). https://doi.org/10.1111/j.1475-4762.2008.875_5.x.
328. Li, T., Chen, Y.-Z., Han, L.-J., et al. (2021). Shortened duration and reduced area of frozen soil in the northern hemisphere. Innovation ***2***(3): 100146. https://doi.org/10.1016/j.xinn.2021.100146.
329. Taylor, L.S., Quincey, D.J., Smith, M.W., et al. (2021). Remote sensing of the mountain cryosphere: Current capabilities and future opportunities for research. Prog. Phys. Geogr. Earth Environ. ***45***(6): 931–964. https://doi.org/10.1177/03091333211023690.
330. Zhang, J., Jia, L., Menenti, M., et al. (2019). Glacier facies mapping using a machine-learning algorithm: The Parlung Zangbo Basin case study. Remote Sens ***11***(4): 452. https://doi.org/10.3390/rs11040452.

331. Alifu, H., Vuillaume, J.-F., Johnson, B.A., et al. (2020). Machine-learning classification of debris-covered glaciers using a combination of Sentinel-1/-2 (SAR/optical), Landsat 8 (thermal) and digital elevation data. Geomorphology ***369***: 107365. https://doi.org/10.1016/j.geomorph.2020.107365.
332. Pastick, N.J., Jorgenson, M.T., Wylie, B.K., et al. (2015). Distribution of near-surface permafrost in Alaska: Estimates of present and future conditions. Remote Sens. Environ. ***168***: 301–315. https://doi.org/10.1016/j.rse.2015.07.019.
333. Kuter, S. (2021). Completing the machine learning saga in fractional snow cover estimation from MODIS Terra reflectance data: Random forests versus support vector regression. Remote Sens. Environ. ***255***: 112294. https://doi.org/10.1016/j.rse.2021.112294.
334. Boulze, H., Korosov, A., and Brajard, J. (2020). Classification of sea ice types in Sentinel-1 SAR data using convolutional neural networks. Remote Sens ***12***(13): 2165. https://doi.org/10.3390/rs12132165.
335. Barbat, M.M., Wesche, C., Werhli, A.V., et al. (2019). An adaptive machine learning approach to improve automatic iceberg detection from SAR images. ISPRS J. Photogramm. Remote Sens. ***156***: 247–259. https://doi.org/10.1016/j.isprsjprs.2019.08.015.
336. Tsai, Y.-L.S., Dietz, A., Oppelt, N., et al. (2019). Wet and dry snow detection using Sentinel-1 SAR data for mountainous areas with a machine learning technique. Remote Sens ***11***(8): 895. https://doi.org/10.3390/rs11080895.
337. Thüring, T., Schoch, M., van Herwijnen, A., et al. (2015). Robust snow avalanche detection using supervised machine learning with infrasonic sensor arrays. Cold Reg. Sci. Technol. ***111***: 60–66. https://doi.org/10.1016/j.coldregions.2014.12.014.
338. Mohajerani, Y., Wood, M., Velicogna, I., et al. (2019). Detection of glacier calving margins with convolutional neural networks: A case study. Remote Sens ***11***(1): 74. https://doi.org/10.3390/rs11010074.
339. Cheng, D., Hayes, W., Larour, E., et al. (2021). Calving front machine (CALFIN): Glacial termini dataset and automated deep learning extraction method for Greenland, 1972-2019. Cryosphere ***15***(3): 1663–1675. https://doi.org/10.5194/tc-15-1663-2021.
340. Qayyum, N., Ghuffar, S., Ahmad, H.M., et al. (2020). Glacial lakes mapping using multi-satellite PlanetScope imagery and deep learning. ISPRS Int. J. Geo-Inf. ***9***(10): 560. https://doi.org/10.3390/ijgi9100560.
341. Huang, L., Luo, J., Lin, Z., et al. (2020). Using deep learning to map retrogressive thaw slumps in the Beiluhe region (Tibetan Plateau) from CubeSat images. Remote Sens. Environ. ***237***: 111534. https://doi.org/10.1016/j.rse.2019.111534.
342. Abolt, C.J., Young, M.H., Atchley, A.L., et al. (2019). Brief communication: Rapid machine-learning-based extraction and measurement of ice wedge polygons in high-resolution digital elevation models. Cryosphere ***13***(1): 237–245.
343. Baumhoer, C.A., Dietz, A.J., Kneisel, C., et al. (2019). Automated extraction of antarctic glacier and ice shelf fronts from Sentinel-1 imagery using deep learning. Remote Sens ***11***(21): 2529. https://doi.org/10.3390/rs11212529.
344. Xiao, H., Zhang, F., He, Q., et al. (2019). Classification of ice crystal habits observed from airborne cloud particle imager by deep transfer learning. Earth Space Sci. ***6***(10): 1877–1886. https://doi.org/10.1029/2019ea000636.
345. Yang, J.W., Jiang, L.M., Lemmetyinen, J., et al. (2021). Improving snow depth estimation by coupling hut-optimized effective snow grain size parameters with the random forest approach. Remote Sens. Environ. ***264***: 112630. https://doi.org/10.1016/j.rse.2021.112630.
346. Broxton, P.D., Van Leeuwen, W.J.D., and Biederman, J.A. (2019). Improving snow water equivalent maps with machine learning of snow survey and lidar measurements. Water Resour. Res. ***55***(5): 3739–3757. https://doi.org/10.1029/2018wr024146.
347. Gao, S., Zeng, J., Li, Z., et al. (2024). Measuring global snow water equivalent from passive microwave remote sensing: opportunities and challenges. The Innovation Geoscience ***2***(2): 100062. https://doi.org/10.59717/j.xinn-geo.2024.100062.
348. Mastro, P., Masiello, G., Serio, C., et al. (2022). Combined IASI-NG and MWS observations for the retrieval of cloud liquid and ice water path: a deep learning artificial intelligence approach. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. ***15***: 3313–3322. https://doi.org/10.1109/jstars.2022.3166992.
349. Guleryuz, D. (2022). Estimation of soil temperatures with machine learning algorithms—Giresun and Bayburt stations in Turkey. Theor. Appl. Climatol. ***147***(1-2): 109–125. https://doi.org/10.1007/s00704-021-03819-2.
350. Guidicelli, M., Huss, M., Gabella, M., et al. (2023). Spatio-temporal reconstruction of winter glacier mass balance in the Alps, Scandinavia, Central Asia and Western Canada (1981-2019) using climate reanalyses and machine learning. Cryosphere ***17***(2): 977–1002. https://doi.org/10.5194/tc-17-977-2023.
351. Erickson, M.L., Elliott, S.M., Brown, C.J., et al. (2021). Machine-learning predictions of high arsenic and high manganese at drinking water depths of the glacial aquifer system, northern continental United States. Environ. Sci. Technol. ***55***(9): 5791–5805. https://doi.org/10.1021/acs.est.0c06740.
352. Ni, J., Wu, T., Zhu, X., et al. (2021). Simulation of the present and future projection of permafrost on the Qinghai-Tibet Plateau with statistical and machine learning models. J. Geophys. Res. Atmos. ***126***(2): e2020JD033402. https://doi.org/10.1029/2020jd033402.
353. Choubin, B., Borji, M., Mosavi, A., et al. (2019). Snow avalanche hazard prediction using machine learning methods. J. Hydrol. X. ***577***: 123929. https://doi.org/10.1016/j.jhydrol.2019.123929.
354. Andersson, T.R., Hosking, J.S., Pérez-Ortiz, M., et al. (2021). Seasonal Arctic sea ice forecasting with probabilistic deep learning. Nat. Commun. ***12***(1): 5124. https://doi.org/10.1038/s41467-021-25257-4.
355. Lee, S., Im, J., Kim, J., et al. (2016). Arctic sea ice thickness estimation from CryoSat-2 satellite data using machine learning-based lead detection. Remote Sens ***8***(9): 698. https://doi.org/10.3390/rs8090698.
356. Cai, Y., Hu, S., Lang, S., et al. (2020). End-to-end classification network for ice sheet subsurface targets in radar imagery. Appl. Sci. ***10***(7): 2501. https://doi.org/10.3390/app10072501.
357. Perry, G.L.W., Seidl, R., Bellvé, A.M., et al. (2022). An outlook for deep learning in ecosystem science. Ecosystems ***25***(8): 1700–1718. https://doi.org/10.1007/s10021-022-00789-y.
358. Christin, S., Hervet, É., and Lecomte, N. (2019). Applications for deep learning in ecology. Methods Ecol. Evol. ***10***(10): 1632–1644. https://doi.org/10.1111/2041-210x.13256.
359. Wang, M., and Sun, B. (2023). Unlocking the connection: Aging as a lens to examine the effects of climate warming. Innovation Life ***1***(1): 100003. https://doi.org/10.59717/j.xinn-life.2023.100003.
360. Huang, J., Zhang, X., Xin, Q., et al. (2019). Automatic building extraction from high-resolution aerial images and LiDAR data using gated residual refinement network. ISPRS J. Photogramm. Remote Sens. ***151***: 91–105. https://doi.org/10.1016/j.isprsjprs.2019.02.019.
361. Liu, S., Zhang, J., Li, J., et al. (2021). Simulating and mitigating extreme urban heat island effects in a factory area based on machine learning. Build. Environ. ***202***: 108051. https://doi.org/10.1016/j.buildenv.2021.108051.
362. Zhang, X., Liu, J., Zeng, J., et al. (2024). Impact of drought-induced forest mortality on terrestrial carbon cycle from remote sensing perspective. Innovation Geosci. ***2***(1): 100057. https://doi.org/10.59717/j.xinn-geo.2024.100057.
363. Rivera Caicedo, J.P., Verrelst, J., Muñoz-Marí, J., et al. (2014). Toward a semiautomatic machine learning retrieval of biophysical parameters. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. ***7***(4): 1249–1259. https://doi.org/10.1109/jstars.2014.2298752.
364. Bao, Z., Zhang, J., Wang, G., et al. (2021). The sensitivity of vegetation cover to climate change in multiple climatic zones using machine learning algorithms. Ecol. Indic. ***124***: 107443. https://doi.org/10.1016/j.ecolind.2021.107443.
365. Verrelst, J., Berger, K., and Rivera-Caicedo, J.P. (2020). Intelligent sampling for vegetation nitrogen mapping based on hybrid machine learning algorithms. IEEE Geosci. Remote Sens. Lett. ***18***(12): 2038–2042. https://doi.org/10.1109/lgrs.2020.3014676.
366. Gessner, U., Machwitz, M., Conrad, C., et al. (2013). Estimating the fractional cover of growth forms and bare surface in savannas: A multi-resolution approach based on regression tree ensembles. Remote Sens. Environ. ***129***: 90–102. https://doi.org/10.1016/j.rse.2012.10.026.
367. Jia, M., Wang, Z., Luo, L., et al. (2023). Global status of mangrove forests in resisting cyclone and tsunami. Innovation Geosci. ***1***(2): 100024. https://doi.org/10.59717/j.xinn-geo.2023.100024.
368. Mustafa, Y.T., Van Laake, P.E., and Stein, A. (2010). Bayesian network modeling for improving forest growth estimates. IEEE Trans. Geosci. Remote Sens. ***49***(2): 639–649. https://doi.org/10.1109/tgrs.2010.2058581.
369. Chen, Q., Zheng, B., Chenu, K., et al. (2022). Unsupervised plot-scale LAI phenotyping via UAV-based imaging, modelling, and machine learning. Plant Phenomics. https://doi.org/10.34133/2022/9768253.
370. Shang, R., Chen, J.M., Xu, M., et al. (2023). China's current forest age structure will lead to weakened carbon sinks in the near future. Innovation ***4***(6): 100515.
371. Fei, Y., Sun, J., Fang, H., et al. (2012). Comparison of different methods for corn LAI estimation over northeastern China. Int. J. Appl. Earth Obs. Geoinf. ***18***: 462–471. https://doi.org/10.1016/j.jag.2011.09.004.
372. Malenovský, Z., Homolová, L., Zurita-Milla, R., et al. (2013). Retrieval of spruce leaf chlorophyll content from airborne image data using continuum removal and radiative transfer. Remote Sens. Environ. ***131***: 85–102. https://doi.org/10.1016/j.rse.2012.12.015.
373. Wang, X., Xiao, X., Zhang, C., et al. (2023). Effects of the 2022 extreme droughts on avian influenza transmission risk in Poyang Lake. Innov. Life ***1***(3): 100044. https://doi.org/10.59717/j.xinn-life.2023.100044.
374. Schiller, C., Schmidtlein, S., Boonman, C., et al. (2021). Deep learning and citizen science enable automated plant trait predictions from photographs. Sci. Rep. ***11***(1): 16395. https://doi.org/10.1038/s41598-021-95616-0.
375. Lang, N., Jetz, W., Schindler, K., et al. (2023). A high-resolution canopy height model of the Earth. Nat. Ecol. Evol. ***7***: 1778–1789. https://doi.org/10.1038/s41559-023-02206-6.
376. Zhang, H., Zhang, H., Xu, K., et al. (2023). A novel framework for stratified-coupled BLS tree trunk detection and DBH estimation in forests (BSTDF) using deep learning and optimization adaptive algorithm. Remote Sens ***15***(14): 3480. https://doi.org/10.3390/rs15143480.
377. Brandt, M., Tucker, C.J., Kariryaa, A., et al. (2020). An unexpectedly large count of trees in the West African Sahara and Sahel. Nature ***587***(7832): 78–82. https://doi.org/10.1038/s41586-020-2824-5.
378. Kälin, U., Lang, N., Hug, C., et al. (2019). Defoliation estimation of forest trees from ground-level images. Remote Sens. Environ. ***223***: 143–153. https://doi.org/10.1101/441733.
379. Persello, C., Tolpekin, V.A., Bergado, J.R., et al. (2019). Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. Remote Sens. Environ. ***231***: 111253. https://doi.org/10.1016/j.rse.2019.111253.
380. Olsson, O., Karlsson, M., Persson, A.S., et al. (2021). Efficient, automated and robust pollen analysis using deep learning. Methods Ecol. Evol. ***12***(5): 850–862. https://doi.org/10.1111/2041-210x.13575.
381. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature ***521***(7553): 436–444. https://doi.org/10.1017/9781108955652.016.

382. Xu, G., Zhu, X., Fu, D., et al. (2017). Automatic land cover classification of geotagged field photos by deep learning. Environ. Model. Softw. ***91***: 127–134. https://doi.org/10.1016/j.envsoft.2017.02.004.
383. Mo, J., Lan, Y., Yang, D., et al. (2021). Deep learning-based instance segmentation method of litchi canopy from UAV-acquired images. Remote Sens ***13***(19): 3919. https://doi.org/10.3390/rs13193919.
384. Wang, C., Hazen, R.M., Cheng, Q., et al. (2021). The deep-time digital earth program: Data-driven discovery in geosciences. Natl. Sci. Rev. ***8***(9): nwab027. https://doi.org/10.1093/nsr/nwab027.
385. Kussul, N., Lavreniuk, M., Skakun, S., et al. (2017). Deep learning classification of land cover and crop types using remote sensing data. IEEE Geosci. Remote Sens. ***14***(5): 778–782. https://doi.org/10.1109/lgrs.2017.2681128.
386. Yang, X., Li, X., Ye, Y., et al. (2019). Road detection and centerline extraction via deep recurrent convolutional neural network U-Net. IEEE Trans. Geosci. Remote Sens. ***57***(9): 7209–7220. https://doi.org/10.1109/tgrs.2019.2912301.
387. Barbosa Martins, G., Cue La Rosa, L.E., Nigri Happ, P., et al. (2021). Deep learning-based tree species mapping in a highly diverse tropical urban setting. Urban For Urban Gree ***64***: 127241. https://doi.org/10.1016/j.ufug.2021.127241.
388. Reiner, F., Brandt, M., Tong, X., et al. (2023). More than one quarter of Africa's tree cover is found outside areas previously classified as forest. Nat. Commun. ***14***(1): 2258. https://doi.org/10.1038/s41467-023-37880-4.
389. Wang, F., Harindintwali, J.D., Yuan, Z., et al. (2021). Technologies and perspectives for achieving carbon neutrality. Innovation ***2***(4): 100180. https://doi.org/10.1016/j.xinn.2021.100180.
390. Jung, M., Schwalm, C., Migliavacca, M., et al. (2020). Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the fluxcom approach. BG ***17***(5): 1343–1365. https://doi.org/10.5194/bg-17-1343-2020.
391. Huang, N., Wang, L., Song, X.-P., et al. (2020). Spatial and temporal variations in global soil respiration and their relationships with climate and land cover. Sci. Adv. ***6***(41): eabb8508. https://doi.org/10.1126/sciadv.abb8508.
392. Richardson, J.G., and Erdelen, W.R. (2021). 2030 is tomorrow: Transformative change for a mistreated mother earth. Foresight ***23***(3): 257–272. https://doi.org/10.1108/fs-03-2020-0029.
393. Ameen, N., Tarhini, A., Reppel, A., et al. (2021). Customer experiences in the age of artificial intelligence. Comput. Hum. Behav. ***114***: 106548. https://doi.org/10.1016/j.chb.2020.106548.
394. Steyn, W.J., and Broekman, A. (2022). Development of a digital twin of a local road network: A case study. J. Test. Eval. ***50***(6): 2901–2915. https://doi.org/10.1520/jte20210043.
395. Samala, N., Katkam, B.S., Bellamkonda, R.S., et al. (2020). Impact of AI and robotics in the tourism sector: A critical insight. J. Tour. Futures ***8***(1): 73–87. https://doi.org/10.1108/jtf-07-2019-0065.
396. Sun, Z., Zhu, M., Zhang, Z., et al. (2021). Artificial intelligence of things (AIoT) enabled virtual shop applications using self-powered sensor enhanced soft robotic manipulator. Adv. Sci. ***8***(14): 2100230. https://doi.org/10.1002/advs.202100230.
397. Ren, T., Xie, Y., and Jiang, L. (2020). Cooperative highway work zone merge control based on reinforcement learning in a connected and automated environment. Transp. Res. Rec. ***2674***(10): 363–374. https://doi.org/10.1177/0361198120935873.
398. Heestermans Svendsen, D., Martino, L., Campos-Taberner, M., et al. (2018). Joint Gaussian Processes for Biophysical Parameter Retrieval. IEEE Trans. Geosci. Remote Sens. ***56***(3): 1718–1727. https://doi.org/10.1109/tgrs.2017.2767205.
399. Li, S., and Zhang, Z. (2024). Prospects for direct air capture. Innovation Energy ***1***(1): 100010. https://doi.org/10.59717/j.xinn-energy.2024.100010.
400. Lowenstern, J.B., Wallace, K., Barsotti, S., et al. (2022). Guidelines for volcano-observatory operations during crises: Recommendations from the 2019 volcano observatory best practices meeting. J. Appl. Volcanol. ***11***(1): 3. https://doi.org/10.1186/s13617-021-00112-9.
401. Sohail, M.T., Hussan, A., Ehsan, M., et al. (2022). Groundwater budgeting of Nari and Gaj formations and groundwater mapping of Karachi, Pakistan. Appl. Water Sci. ***12***(12): 267. https://doi.org/10.1007/s13201-022-01795-0.
402. Bruzzone, O.A., Perri, D.V., and Easdale, M.H. (2023). Vegetation responses to variations in climate: A combined ordinary differential equation and sequential Monte Carlo estimation approach. Ecol. Inf. ***73***: 101913. https://doi.org/10.1016/j.ecoinf.2022.101913.
403. Resolution, A. (2015). Res/70/1 Transforming our world: The 2030 agenda for sustainable development. 70th UNGA ***25***: 86–97. https://doi.org/10.5040/9781509934058.0025.
404. Jean, N., Burke, M., Xie, M., et al. (2016). Combining satellite imagery and machine learning to predict poverty. Science ***353***(6301): 790–794. https://doi.org/10.1126/science.aaf7894.
405. Guo, H., Huang, L., and Liang, D. (2022). Further promotion of sustainable development goals using science, technology, and innovation. Innovation. ***3***(6). https://doi.org/10.1016/j.xinn.2022.100325.
406. Vinuesa, R., Azizpour, H., Leite, I., et al. (2020). The role of artificial intelligence in achieving the sustainable development goals. Nat. Commun. ***11***(1): 233–310. https://doi.org/10.1038/s41467-019-14108-y.
407. Gupta, S., Langhans, S.D., Domisch, S., et al. (2021). Assessing whether artificial intelligence is an enabler or an inhibitor of sustainability at indicator level. J Transp Eng ***4***: 100064. https://doi.org/10.1016/j.treng.2021.100064.
408. Nasir, O., Javed, R.T., Gupta, S., et al. (2023). Artificial intelligence and sustainable development goals nexus via four vantage points. Technol. Soc. ***72***: 102171. https://doi.org/10.1016/j.techsoc.2022.102171.
409. Acemoglu, D., and Restrepo, P. (2018). Artificial intelligence, automation, and work (Econ. of AI: An Agenda), pp. 197–236. https://doi.org/10.7208/chicago/9780226613475.003.0008.
410. Brynjolfsson, E., and McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies.
411. Guo, H., Liang, D., Sun, Z., et al. (2022). Measuring and evaluating SDG indicators with big Earth data. Sci. Bull. ***67***(17): 1792–1801. https://doi.org/10.1016/j.scib.2022.07.015.
412. Cheng, Y., Li, C., Xu, Y., et al. (2024). Extreme impacts on electric power systems from non-catastrophic meteorological conditions. Innovation Energy ***1***(1): 100008. https://doi.org/10.59717/j.xinn-energy.2024.100008.
413. Chen, X. (2024). Green and low-carbon energy-use. Innovation Energy ***1***(1): 100003. https://doi.org/10.59717/j.xinn-energy.2024.100003.
414. Deng, Y., Baeyens, J., Liu, H., et al. (2024). Renewable electricity and 'green' feedstock-based chemicals will foster industrial sustainability. Innovation Energy ***1***(2): 100016. https://doi.org/10.59717/j.xinn-energy.2024.100016.
415. Shen, M., Duan, H., Cao, Z., et al. (2020). Sentinel-3 OLCI observations of water clarity in large lakes in eastern China: Implications for SDG 6.3.2 evaluation. Remote Sens. Environ. ***247***: 111950. https://doi.org/10.1016/j.rse.2020.111950.
416. Luo, L., Wang, X., Guo, H., et al. (2022). Eighteen years (2001-2018) of forest habitat loss across the Asian elephant's range and its drivers. Sci. Bull. ***67***: 1513–1516. https://doi.org/10.1016/j.scib.2022.04.013.
417. Luo, J., Wu, M., and Yang, Y. (2024). Unlocking a 30 billion market opportunity with carbon dioxide utilization. Innovation ***1***(1): 100009. https://doi.org/10.59717/j.xinn-energy.2024.100009.
418. Li, C., Hong, D., Zhang, B., et al. (2024). Interpretable foundation model as decryptor peering into Earth system. Innovation: 2666–6758. https://doi.org/10.1016/j.xinn.2024.100682.
419. Khan Raiaan, M.A., Hossain Mukta, M.S., Fatema, K., et al. (2023). A review on large language models: Architectures, applications, taxonomies, open issues, and challenges. IEEE Acc ***12***: 26839–26874. https://doi.org/10.1109/ACCESS.2024.3365742.
420. Van Dis, E.A.M., Bollen, J., Zuidema, W., et al. (2023). ChatGPT: Five priorities for research. Nature ***614***(7947): 224–226. https://doi.org/10.1038/d41586-023-00288-7.
421. Sun, X., Yang, Y., Jia, H., et al. (2022). Physics-aware training for the physical machine learning model building. Innovation ***3***(5): 100287. https://doi.org/10.1016/j.xinn.2022.100287.
422. Xu, Y., Wang, F., An, Z., et al. (2023). Artificial intelligence for science—bridging data to wisdom. Innovation ***4***(6): 100525. https://doi.org/10.1016/j.xinn.2023.100525.
423. Foroumandi, E., Moradkhani, H., Sanchez-Vila, X., et al. (2023). ChatGPT in hydrology and Earth sciences: Opportunities, prospects, and concerns. Water Resour. Res. ***59***(10): e2023WR036288. https://doi.org/10.1029/2023wr036288.
424. Yuan, X., and Han, S. (2021). Single-pixel neutron imaging with artificial intelligence: Breaking the barrier in multi-parameter imaging, sensitivity, and spatial resolution. Innovation ***2***(2): 100100. https://doi.org/10.1016/j.xinn.2021.100100.
425. Li, C., Zhang, B., Hong, D., et al. (2024). Casformer: Cascaded transformers for fusion-aware computational hyperspectral imaging. Inform. Fusion ***108***: 102408. https://doi.org/10.1016/j.inffus.2024.102408.
426. Xu, Q., Shi, Y., Guo, J., et al. (2023). UCDformer: Unsupervised change detection using a transformer-driven image translation. IEEE Trans. Geosci. Remote Sens. ***61***: 1–7. https://doi.org/10.1109/tgrs.2023.3305334.
427. Cha, K., Seo, J., and Lee, T. (2024). A billion-scale foundation model for remote sensing images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2151–2217. https://doi.org/10.1109/jstars.2024.3401772.
428. Nguyen, T., Brandstetter, J., Kapoor, A., et al. (2023). Climax: A foundation model for weather and climate. Preprint at arXiv. https://doi.org/10.48550/arXiv.2301.10343.
429. Bi, Z., Zhang, N., Xue, Y., et al. (2024). Oceangpt: A large language model for ocean science tasks. Preprint at arXiv. https://doi.org/10.48550/arXiv.2310.02031.
430. Li, C., Zhang, B., Hong, D., et al. (2023). LRR-Net: An interpretable deep unfolding network for hyperspectral anomaly detection. IEEE Trans. Geosci. Remote Sens. ***61***: 1–12. https://doi.org/10.1109/tgrs.2023.3279834.
431. Li, C., Zhang, B., Hong, D., et al. (2024). Learning disentangled priors for hyperspectral anomaly detection: A coupling model-driven and data-driven paradigm. IEEE Trans. Neural Netw. Learn. Syst. ***4***. https://doi.org/10.1109/tnnls.2024.3401589.
432. Yao, J., Hong, D., Li, C., et al. (2024). SpectralMamba: Efficient Mamba for hyperspectral image classification. Preprint at arXiv ***4***. https://doi.org/10.48550/arXiv.2404.08489.
433. Hong, D., Zhang, B., Li, X., et al. (2024). SpectralGPT: Spectral remote sensing foundation model. IEEE T. Pattern Anal. ***46***(8): 5227–5244. https://doi.org/10.1109/tpami.2024.3362475.
434. Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. Preprint at arXiv. https://doi.org/10.48550/arXiv.2010.11929.
435. Kirillov, A., Mintun, E., Ravi, N., et al. (2023). Segment Anything (ICCV), pp. 4015–4026. https://doi.org/10.1109/iccv51070.2023.00371.
436. IBM and NASA (2023). An Open-Source, Geospatial Foundational Model.
437. Zhang, W., Cai, M., Zhang, T., et al. (2024). EarthGPT: A universal multi-modal large language model for multi-sensor image comprehension in remote sensing domain. IEEE T. Geosci. Remote Sens. ***62***: 1–20.
438. Pathak, J., Subramanian, S., Harrington, P., et al. (2022). FourCastNet: A global data-driven high-resolution weather model using adaptive Fourier neural operators. Preprint at arXiv. https://doi.org/10.48550/arXiv.2202.11214.
439. Andrychowicz, M., Espeholt, L., Li, D., et al. (2023). Deep learning for day forecasts from sparse observations. Preprint at arXiv. https://doi.org/10.48550/arXiv.2306.06079.

440. Price, I., Sanchez-Gonzalez, A., Alet, F., et al. (2023). GenCast: Diffusion-based ensemble forecasting for medium-range weather. Preprint at arXiv. https://doi.org/10.48550/arXiv.2312.15796.
441. Thulke, D., Gao, Y., Pelser, P., et al. (2024). ClimateGPT: Towards AI synthesizing interdisciplinary research on climate change. Preprint at arXiv. https://doi.org/10.48550/arXiv.2401.09646.
442. Goecks, V.G., and Waytowich, N.R. (2023). DisasterResponseGPT: Large language models for accelerated plan of action development in disaster response scenarios. Preprint at arXiv. https://doi.org/10.48550/ARXIV.2306.17271.
443. Darlow, L.N., Deng, Q., Hassan, A., et al. (2024). DAM: Towards a Foundation Model for Forecasting (ICLR). https://doi.org/10.1037/e518442013-089.
444. Zhang, T., Gao, P., Dong, H., et al. (2022). Consecutive pre-training: A knowledge transfer learning strategy with relevant unlabeled data for remote sensing domain. Remote Sens ***14***(22): 5675. https://doi.org/10.3390/rs14225675.
445. Manas, O., Lacoste, A., Giro-i Nieto, X., et al. (2021). Seasonal Contrast: Unsupervised Pre-training from Uncurated Remote Sensing Data (ICCV), pp. 9414–9423. https://doi.org/10.1109/iccv48922.2021.00928.
446. Wang, D., Zhang, Q., Xu, Y., et al. (2022). Advancing plain vision transformer toward remote sensing foundation model. IEEE Trans. Geosci. Remote Sens. ***61***: 1–15. https://doi.org/10.1109/tgrs.2022.3222818.
447. Cong, Y., Khanna, S., Meng, C., et al. (2022). SatMAE: Pre-training transformers for temporal and multi-spectral satellite imagery. NeurIPS ***35***: 197–211. https://doi.org/10.48550/ARXIV.2207.08051.
448. Jain, U., Wilson, A., and Gulshan, V. (2022). Multimodal contrastive learning for remote sensing tasks. Preprint at arXiv. https://doi.org/10.48550/arXiv.2209.02329.
449. Li, W., Chen, H., and Shi, Z. (2021). Semantic segmentation of remote sensing images with self-supervised multitask representation learning. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. ***14***: 6438–6450. https://doi.org/10.1109/jstars.2021.3090418.
450. Reed, C.J., Gupta, R., Li, S., et al. (2023). Scale-MAE: A scale-aware masked autoencoder for multiscale geospatial representation learning. ICCV: 408–4099. https://doi.org/10.1109/iccv51070.2023.00378.
451. Hong, D., Zhang, B., Li, H., et al. (2023). Cross-city matters: A multimodal remote sensing benchmark dataset for cross-city semantic segmentation using high-resolution domain adaptation networks. Remote Sens. Environ. ***299***: 113856. https://doi.org/10.1016/j.rse.2023.113856.
452. Hong, D., Yao, J., Li, C., et al. (2023). Decoupled-and-coupled networks: Self-supervised hyperspectral image super-resolution with subpixel fusion. IEEE Trans. Geosci. Remote Sens. ***61***: 1–12. https://doi.org/10.1109/tgrs.2023.3324497.
453. Hong, D., Li, C., Zhang, B., et al. (2024). Multimodal artificial intelligence foundation models: Unleashing the power of remote sensing big data in earth observation. The Innovation Geoscience ***2***(1): 100055. https://doi.org/10.59717/j.xinn-geo.2024.100055.
454. Mendieta, M., Han, B., Shi, X., et al. (2023). Towards Geospatial Foundation Models via Continual Pretraining. ICCV: 16760–16770. https://doi.org/10.1109/iccv51070.2023.01541.
455. Fuller, A., Millard, K., and Green, J.R. (2022). SatViT: Pretraining transformers for earth observation. IEEE Geosci. Remote Sens. ***19***: 1–5. https://doi.org/10.1109/lgrs.2022.3201489.
456. Sun, X., Wang, P., Lu, W., et al. (2022). RingMo: A remote sensing foundation model with masked image modeling. IEEE Trans. Geosci. Remote Sens. ***61***: 1–22. https://doi.org/10.1109/tgrs.2022.3194732.
457. Scheibenreif, L., Mommert, M., and Borth, D. (2023). Masked vision transformers for hyperspectral image classification. CVPR: 2165–2175. https://doi.org/10.1109/cvprw59228.2023.00210.
458. Gao, Z., Shi, X., Wang, H., et al. (2022). EarthFormer: Exploring space-time transformers for earth system forecasting. NeurIPS ***35***: 25390–25403. https://doi.org/10.1109/lapc.2009.5352582.
459. Deng, C., Zhang, T., He, Z., et al. (2023). Learning a foundation language model for geoscience knowledge understanding and utilization. Preprint at arXiv. https://doi.org/10.1145/3616855.3635772.
460. Ma, C., Morrison, S.M., Muscente, A.D., et al. (2022). Incorporate temporal topology in a deep-time knowledge base to facilitate data-driven discovery in geoscience. Geosci. Data J. ***10***(4): 489–499. https://doi.org/10.1002/gdj3.171.
461. Normile, D. (2019). Earth scientists plan a 'geological google. Science ***363***: 917. https://doi.org/10.1126/science.363.6430.917.
462. Fan, J.-x., Shen, S.-z., Erwin, D.H., et al. (2020). A high-resolution summary of Cambrian to early Triassic marine invertebrate biodiversity. Science ***367***(6475): 272–277. https://doi.org/10.1126/science.aax4953.
463. Muscente, A.D., Bykova, N., Boag, T.H., et al. (2019). Ediacaran biozones identified with network analysis provide evidence for pulsed extinctions of early complex life. Nat. Commun. ***10***(1): 911. https://doi.org/10.1038/s41467-019-08837-3.
464. Peters, S.E., Husson, J.M., and Wilcots, J. (2017). The rise and fall of stromatolites in shallow marine environments. Geology ***45***(6): 487–490. https://doi.org/10.1130/g38931.1.
465. Hazen, R.M., Downs, R.T., Eleish, A., et al. (2019). Data-driven discovery in mineralogy: Recent advances in data resources, analysis, and visualization. Engineering ***5***(3): 397–405. https://doi.org/10.1016/j.eng.2019.03.006.
466. Peters, S.E., and Husson, J.M. (2017). Sediment cycling on continental and oceanic crust. Geology ***45***(4): 323–326. https://doi.org/10.1130/g38861.1.
467. Liu, C., Knoll, A.H., and Hazen, R.M. (2017). Geochemical and mineralogical evidence that Rodinian assembly was unique. Nat. Commun. ***8***(1): 1950. https://doi.org/10.1038/s41467-017-02095-x.
468. Wright, N., Zahirovic, S., Muller, R.D., et al. (2013). Towards community-driven paleogeographic reconstructions: Integrating open-access paleogeographic and paleobiology data with plate tectonics. Biogeosciences ***10***(3): 1529–1541. https://doi.org/10.5194/bg-10-1529-2013.
469. Muller, R.D., and Dutkiewicz, A. (2018). Oceanic crustal carbon cycle drives 26-million-year atmospheric carbon dioxide periodicities. Sci. Adv. ***4***(2): eaaq0500. https://doi.org/10.1126/sciadv.aaq0500.
470. Mallard, C., Coltice, N., Seton, M., et al. (2016). Subduction controls the distribution and fragmentation of Earth's tectonic plates. Nature ***535***(7610): 140–143. https://doi.org/10.1038/nature17992.
471. Wang, C., Wang, T., Chen, X., et al. (2017). Paleoclimate implications for future climate change. Earth Sci. Front. ***24***(1): 1–17. https://doi.org/10.1007/978-3-642-79066-9_26.
472. Zhang, L., Wang, C., Li, X., et al. (2016). A new paleoclimate classification for deep time. Palaeogeogr. Palaeoclimatol. Palaeoecol. ***443***: 98–106. https://doi.org/10.1016/j.palaeo.2015.11.041.
473. Perez-Ortiz, M., Durán-Rosal, A.M., Gutiérrez, P.A., et al. (2019). On the use of evolutionary time series analysis for segmenting paleoclimate data. Neurocomputing ***326***: 3–14. https://doi.org/10.1016/j.neucom.2016.11.101.
474. Hu, X., Xu, Y., Ma, X., et al. (2023). Knowledge system, ontology, and knowledge graph of the deep-time digital Earth (DDE): Progress and perspective. J. Earth Sci. ***34***(5): 1323–1327. https://doi.org/10.1007/s12583-023-1930-1.
475. Zhu, Y., Dai, X., Yang, J., et al. (2023). One-stop sharing and service system for geoscience knowledge graph. Geol. J. China Univ. ***29***(3): 325. https://doi.org/10.1109/ickii.2018.8569149.
476. Yu, C., Zhang, L., Hou, M., et al. (2023). Climate paleogeography knowledge graph and deep time paleoclimate classifications. Geosci. Front. ***14***(5): 101450. https://doi.org/10.1016/j.gsf.2022.101450.
477. Wang, H., Zhong, H., Chen, A., et al. (2023). A knowledge graph for standard carbonate microfacies and its application in the automatical reconstruction of the relative sea-level curve. Geosci. Front. ***14***(5): 101535. https://doi.org/10.1016/j.gsf.2023.101535.
478. Deng, C., Jia, Y., Xu, H., et al. (2021). GAKG: A multimodal geoscience academic knowledge graph. Proc. 30th ACM Int. Conf. Inf. and Knowl. Manag. 4445–4454. https://doi.org/10.1145/3459637.3482003.
479. Dias, P., and Lunga, D. (2022). Embedding ethics and trustworthiness for sustainable AI in Earth sciences: Where do we begin? IGARSS: 4639–4642. https://doi.org/10.1109/igarss46834.2022.
480. Coffer, M.M. (2020). Balancing privacy rights and the production of high-quality satellite imagery. Environ. Sci. Technol. ***54***(11): 6453–6455. https://doi.org/10.1021/acs.est.0c02365.
481. Patino, J.E., Duque, J.C., Pardo-Pascual, J.E., et al. (2014). Using remote sensing to assess the relationship between crime and the urban layout. Appl. Geogr. ***55***: 48–60. https://doi.org/10.1016/j.apgeog.2014.08.016.
482. Xu, Y., Bai, T., Yu, W., et al. (2023). AI security for geoscience and remote sensing: Challenges and future trends. IEEE Geosci. Remote Sens. Mag. ***11***(2): 60–85. https://doi.org/10.1109/mgrs.2023.3272825.
483. Dräger, N., Xu, Y., and Ghamisi, P. (2023). Backdoor attacks for remote sensing data with wavelet transform. IEEE Trans. Geosci. Remote Sens. ***61***: 1–15. https://doi.org/10.1109/tgrs.2023.3289307.
484. Janssen, T.A.J., Jones, M.W., Finney, D., et al. (2023). Extratropical forests increasingly at risk due to lightning fires. Nat. Geosci. ***16***(12): 1136–1144. https://doi.org/10.1038/s41561-023-01322-z.
485. Kraft, B., Jung, M., Körner, M., et al. (2021). Towards hybrid modeling of the global hydrological cycle. Hydrol. Earth Syst. Sci. Discuss. ***2021***: 1–40. https://doi.org/10.5194/hess-26-1579-2022.
486. Wang, C., Jiang, S., Zheng, Y., et al. (2024). Distributed hydrological modeling with physics-encoded deep learning: A general framework and its application in the Amazon. Water Resour. Res. ***60***(4): e2023WR036170. https://doi.org/10.1029/2023wr036170.
487. Acuña Espinoza, E., Loritz, R., Álvarez Chaves, M., et al. (2023). To bucket or not to bucket? Analyzing the performance and interpretability of hybrid hydrological models with dynamic parameterization. EGUsphere: 1–22. https://doi.org/10.5194/hess-28-2705-2024.
488. Vonk, J.E., Tank, S.E., and Walvoord, M.A. (2019). Integrating hydrology and biogeochemistry across frozen landscapes. Nat. Commun. ***10***(1): 5377. https://doi.org/10.1038/s41467-019-13361-5.
489. Hickmon, N.L., Varadharajan, C., Hoffman, F.M., et al. (2022). Artificial intelligence for Earth system predictability (AI4ESP) (2021 workshop report) (Argonne Natl. Lab.). https://doi.org/10.2172/1888810.
490. De Paz, J.F., Bajo, J., González, Á., et al. (2012). Combining case-based reasoning systems and support vector regression to evaluate the atmosphere-ocean interaction. Knowl. Inf. Syst. ***30***: 155–177. https://doi.org/10.1007/s10115.
491. Martinez Amaya, J., Radin, C., Nieves, V., et al. (2023). An AI hybrid predictive tool for extreme hurricane forecasting. Technical report, Copernicus Mtgs. EGU23-12333. https://doi.org/10.5194/egusphere-egu23-12333.
492. Moore, W.B., Simon, J.I., and Webb, A.A.G. (2017). Heat-pipe planets. Earth Planet. Sci. Lett. ***474***: 13–19. https://doi.org/10.1016/j.epsl.2017.06.015.
493. Navarro, T., Merlis, T.M., Cowan, N.B., et al. (2022). Atmospheric gravitational tides of Earth-like planets orbiting low-mass stars. Planet. Sci. J. ***3***(7): 162. https://doi.org/10.3847/psj/ac76cd.
494. Warren-Rhodes, K., Cabrol, N.A., Phillips, M., et al. (2023). Orbit-to-ground framework to decode and predict biosignature patterns in terrestrial analogues. Nat. Astron. ***7***(4): 406–422. https://doi.org/10.1038/s41550-022-01882-x.

495. Cleaves, H.J., Hystad, G., Prabhu, A., et al. (2023). A robust, agnostic molecular biosignature based on machine learning. Proc. Natl. Acad. Sci. USA ***120***(41): e2307149120. https://doi.org/10.1073/pnas.2307149120.
496. Maskey, M., Ramachandran, R., Gurung, I., et al. (2022). Artificial intelligence vis-à-vis data systems (IGARSS), pp. 5081–5084. https://doi.org/10.1109/igarss46834.2022.9883626.
497. Ad Hoc Expert Group (2020). First draft of the recommendation of the ethics of artificial intelligence.
498. Goodman, B., and Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a "right to explanation". AI Mag ***38***(3): 50–57. https://doi.org/10.1609/aimag.v38i3.2741.

ACKNOWLEDGMENTS

This work was partially supported by National Natural Science Foundation of China (T2225019, 41925007, 62372470, U21A2013, 42201415, 42022054, 42241109, 42077156, 52121006, 42090014, and 42325107), the National Key R&D Programme of China (2022YFF0 500), the Youth Innovation Promotion Association CAS (2023112), the Strategic Priority Research Program of CAS (XDA23090303), and the RECLAIM Network Plus (EP/W034034/1).

AUTHOR CONTRIBUTIONS

C.L., Y. Xie, and A.P. wrote the introduction. M.C., F.Z., and Z.Q. wrote the paradigms section. S.W., L.Y., C.Y., W.H., T.S., Z.S., T.Q., and Z.C. wrote the AI-driven geoscience paradigms section. C.S., S.Y., N.L., and Y.Z. wrote the atmosphere section. H.Z. wrote the lithosphere section. J. Zeng, H.S., C.Z., and J. Zhang wrote the hydrosphere section. T.Z. wrote the cryosphere section. L. Wang, N.H., and C.H. wrote the biosphere section. L.L., H.Z., and W.Z. wrote the other domains section. T.Z., H.L., J.S., and D.F. revised the typical cases section. C.O., Q.X., Y.W., S.W., and D.H. wrote the large models in geoscience section. J. Zhang, Z.W., Y.L., and T.Z. wrote the challenges and future perspectives in AI for geoscience section. A.P., Lizhe Wang, Y. Xu, F. Wang, B.Z., P.K., and J.L. revised the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

LEAD CONTACT WEBSITE

https://grzy.cug.edu.cn/LizheWang.