

ISSN 2096-742X
CN 10-1649/TP文献 CSTR:
32002.14.jfdc.
CN10-1649/
TP.2026.02.001文献 DOI:
10.11871/jfdc.issn.
2096-742X.2026.
02.001

页码:3-14

获取全文



专刊:冰冻圈大数据挖掘分析关键技术及应用

Special Issue: Key Technologies and Applications of Cryospheric Big Data Mining and Analysis

冰冻圈“大数据+AI+模型”耦合研究范式探索

张耀南^{1,2*}, 刘景琦^{1,2,3}, 康建芳^{1,2,3}, 南卓铜⁴, 田文彪⁴, 敏玉芳^{1,2}, 赵书萍⁴, 王保得^{5,6}

1. 中国科学院西北生态环境资源研究院,甘肃 兰州 730030

2. 国家冰川冻土沙漠科学数据中心,甘肃 兰州 730030

3. 中国科学院大学,北京 100049

4. 南京师范大学,江苏 南京 210023

5. 中国科学院新疆生态与地理研究所,新疆 乌鲁木齐 830011

6. 干旱区生态安全与可持续发展全国重点实验室,新疆 乌鲁木齐 830011

摘要:【目的】冰冻圈作为地球圈层系统的重要部分,其多要素非线性联动、跨圈层耦合及灾害链对全球气候、水文循环及生态安全具有深刻影响。传统物理模型在多尺度、多要素联动表征冰冻圈变化方面存在局限,亟需建立新的科学研究范式来满足冰冻圈大尺度变化监测、快速演变解析及灾害预警等实际需求。【方法】本文提出了“大数据+AI+模型”冰冻圈研究新范式,研发了支持该范式研究的全球冰冻圈研究引擎(Global Cryosphere research Engine, GCE),形成了数据管理、模型构建、AI算法应用、任务流程编排等全链条一体化平台,具备多源异构数据融合、跨语言模型协同、高性能计算调度与智能分析能力。【结论】基于GCE平台,实现了“物理模型嵌入AI”“AI嵌入物理模型”“数据驱动耦合AI的参数优化”模式的构建,并围绕青藏高原土壤水分模拟试验,建立了“观测数据+深度学习(分层LSTM)+物理模型(Noah-MP)”多要素联动框架,土壤水分分层模拟结果与实测值的平均相关系数超过0.8,显著高于Noah模型。研究表明,物理模型融合人工智能的研究范式可有效刻画冰冻圈过程,验证了GCE平台的可用性,可为生态环境、冰冻圈灾害等新范式研究提供环境支持,推动冰冻圈向跨区域多要素联动的智能化研究深入。

关键词: 冰冻圈;大数据;人工智能;模型;冰冻圈研究引擎GCE

Exploration of An Integrated “Big Data+AI+Modeling” Research Paradigm for Cryosphere Studies

ZHANG Yaonan^{1,2*}, LIU Jingqi^{1,2,3}, KANG Jianfang^{1,2,3}, NAN Zhuotong⁴, TIAN Wenbiao⁴,
MIN Yufang^{1,2}, ZHAO Shuping⁴, Wang Baode^{5,6}1. Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences,
Lanzhou, Gansu 730030, China

基金项目: 国家重点研发计划(2022YFF0711704);2022年度新疆交通运输行业科技项目(2022-ZD-006);新疆交投2021年揭榜挂帅科技项目(ZKXFWCG2022060004);新疆交通设计院公司科研基金(KY2022041101);新疆维吾尔自治区科技支疆项目(2024E02030)

*通信作者: 张耀南(E-mail: yaonan@lzb.ac.cn)

2. National Cryosphere Desert Data Center, Lanzhou, Gansu 730030, China

3. University of Chinese Academy of Sciences, Beijing 100049, China

4. Nanjing Normal University, Nanjing, Jiangsu 210023, China

5. Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Chinese Academy of Sciences,

Urumqi, Xinjiang 830011, China

6. State Key Laboratory of Ecological Security and Sustainable Development in Arid Areas,

Urumqi, Xinjiang 830011, China

Abstract: [Objective] As a vital component of Earth's system, the cryosphere profoundly impacts global climate, hydrological cycles, and ecological security through multi-element nonlinear interactions, cross-sphere couplings, and long disaster chains. Traditional physical models exhibit limitations in characterizing multi-scale cryospheric changes, failing to meet practical needs for large-scale monitoring, rapid evolution analysis, and hazard early-warning. [Methods] This paper proposes a “Big Data+AI+Model” research paradigm for cryospheric studies and develops the Global Cryosphere research Engine (GCE) to support this framework. The GCE forms an integrated platform encompassing data management, model construction, AI algorithm application, and task workflow orchestration, with capabilities for multi-source heterogeneous data fusion, cross-language model collaboration, high-performance computing scheduling, and intelligent analysis. [Conclusions] Based on the GCE platform, model construction approaches such as “embedding AI into physical models”, “embedding physical models into AI”, and “data-driven parameter optimization coupled with AI” have been implemented. Focusing on soil moisture simulation experiments over the Tibetan Plateau, a novel multi-element interaction framework integrating “observation data+deep learning (hierarchical LSTM)+physical model (Noah-MP)” was developed. The average correlation between simulated and observed layered soil moisture exceeded 0.8, significantly outperforming the Noah model. The results demonstrate the feasibility of a research paradigm that combines data-driven approaches, physical model support, and artificial intelligence to effectively characterize cryospheric processes, while also validating the utility of the GCE platform. This framework can provide a supportive environment for new research paradigms in ecological environments and cryospheric disasters, promoting intelligent cryospheric research with cross-regional multi-element interactions.

Keywords: cryosphere; big data; AI; model; GCE (Global Cryosphere research Engine)

引言

冰冻圈作为世界气象组织(WMO)认定的地球气候系统第五大圈层,涵盖冰川、冻土、积雪、冰盖、湖泊与河冰、海冰、冰架、冰山及大气中的冻结水等要素,其动态演变过程深刻影响全球水文循环格局、气候系统稳定性、生态环境安全及社会经济可持续发展^[1-4]。在世界气候研究计划(WCRP)气候与冰冻圈计划(CliC)的推动下,冰冻圈科学已逐渐发展为国际地球系统科学的重要前沿^[5]。

早期冰冻圈研究主要依赖野外调查和传统测绘,聚焦冰川、冻土的分布与特征^[6-9],受限于

观测精度与覆盖度,整体停留在局地、单要素阶段^[10]。随着计算方法与统计分析的引入,数据驱动思维逐渐萌芽^[11]。20世纪末以来,卫星遥感观测的广泛应用,推动了冰冻圈动态模拟与多源耦合^[12-18]研究。近年来,随着大数据技术的持续发展与AI for Science领域的快速推进,人工智能(AI)技术已在冰川边界提取、冻土预测、冰湖监测及海冰识别等任务中成功应用,其精度与效率均优于传统方法^[19-24]。目前,在大数据与人工智能深度融合的背景下,冰冻圈人工智能应用研究正聚焦于四方面:一是依托人工智能开展复杂系统建模,重点解决冰川消融预测、冰盖稳定性评估等难题;二是借助人工智能技术理解冻土与气

候的反馈机制,探究北极及高海拔冻土区碳释放对全球变暖的潜在影响;三是基于数据驱动融合人工智能技术,围绕“冰-洋-大气”相互作用、高精度降水模拟与海平面上升等关联机制,开展多圈层耦合研究;四是构建卫星遥感、无人机与地面观测一体化的多源数据融合网络,形成高质量体系化数据集,进一步提升全球冰冻圈的监测能力。尽管目前研究已取得了一定进展,但在支撑冰冻圈多变量联动综合研究方面仍存在不足,难以全面准确刻画气候变化下冰冻圈的响应过程。大数据驱动下人工智能耦合物理模型构建新的冰冻圈变化过程表征模型,为冰冻圈研究提供了新的契机。基于大模型的变化预测、多尺度物理机制耦合,以及AI驱动的数据同化与补全技术,能够促进数据管理、知识管理与模型开发的深度融合,推动跨学科、跨区域的多要素综合研究,并逐步实现“自然语言驱动的科研全流程”。现有平台 Google Earth Engine (GEE) 虽在遥感数据处理方面具有优势,但在物理过程建模与智能融合方面仍存在明显局限^[25]。

在此背景下,亟需发展冰冻圈新型研究范

式,构建具备大数据管理、智能建模与多源融合能力的冰冻圈研究平台,实现从局地到大尺度、从瞬时过程探测到长期演变模拟的多尺度集成分析,精准支撑冰冻圈大尺度变化监测、快速演变过程解析及灾害预警等核心研究场景,为学科从滞后性分析向前瞻性探索的转型提供核心技术载体。

1 冰冻圈研究研究范式演进

1.1 从观测主导迈向第四范式

冰冻圈研究范式经历了由局地观测到遥感监测,再到多圈层耦合模拟的发展过程(图1)。早期研究依赖地面观测与气象站数据,主要聚焦单一要素的物理特性与统计分析^[26, 27]。进入21世纪,遥感技术与寒区水文模型显著提升了对冰冻圈的高分辨率监测与模拟能力^[28-32]。随着大数据的发展,数据逐渐成为科学探索的核心资源与研究基础,支撑问题识别、模型构建与研究设计。

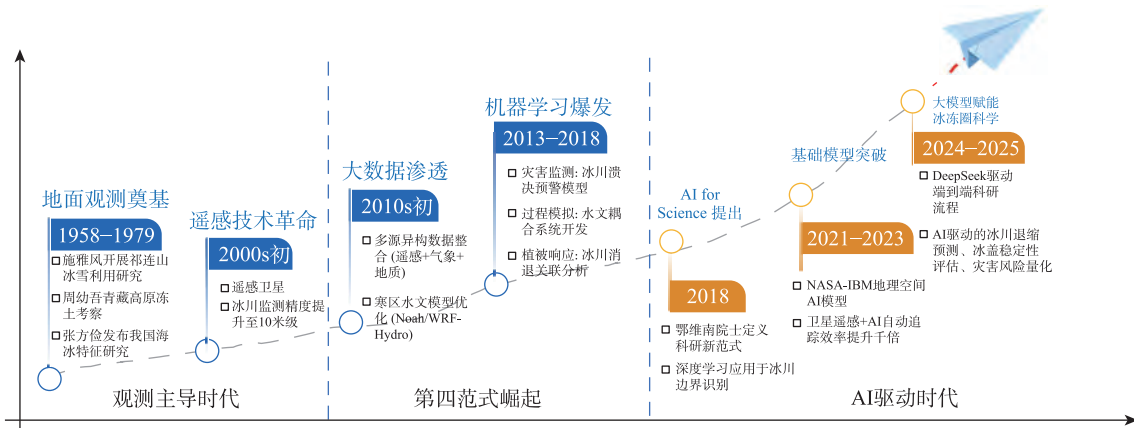


图1 冰冻圈研究发展进程

Fig.1 Progress of Cryosphere research and development

这一演进推动学科研究从经验归纳与理论推演,迈向以数据密集型科学发现(Data-Intensive Scientific Discovery)为核心的第四范式(图1)^[33]。与传统范式相比,数据密集型科学不仅强调数据的建模、存储与分析,更注重其知识价值

的挖掘与复用,从而促进研究由碎片化走向系统化,并实现从定性理解向定量预测的转变^[34]。其核心在于利用统计分析、机器学习等方法,从大规模异构数据中提取潜在规律,形成以数据驱动的协同创新范式。

第四范式在冰冻圈研究中展现出三大特征。一是数据驱动的智能决策,通过整合多源异构数据构建高精度模型,实现冰川消融、海冰变化等过程的实时监测与趋势预测;二是跨学科协同,将冰川、冻土、积雪与地质灾害等领域知识与先进计算方法融合,推动复杂问题的系统化求解;三是方法论创新,通过数据驱动与机器学习的研究思路逐步取代传统假设驱动范式,为冰冻圈要素的感知、过程建模与预测提供新的路径。

目前,第四范式已在冰冻圈研究中展现出显著成效。基于机器学习的研究显著提升了环境与灾害监测能力^[35,36],并为冰川消退下的植被格局演变、冰冻圈水文过程模拟提供了关键支撑^[37-39],标志着该领域已由观测主导迈向数据驱动的新阶段。

1.2 从数据驱动迈向人工智能驱动

2018年,鄂维南院士提出“AI for Science”理念,标志着科研范式开始由数据驱动向人工智能(AI)驱动转变(图1)^[40]。该范式以大规模数据与复杂算法为核心,重塑知识生成路径与科研组织方式,促进跨学科深度融合^[41,42]。近二十年CNKI与Web of Science的计量分析表明,AI相关研究成果呈指数级增长^[43],为范式转型奠定了数据与方法基础。目前,以ChatGPT、DeepSeek为代表的大模型通过技术要素的深度融合,推动人工智能从辅助工具升级为科研合作伙伴,正式开启“人机共智”新阶段^[44]。AI现已覆盖科研全流程,包括问题识别、数据处理、实验设计与结果验证,并在分子结构识别、反应路径预测及药物递送优化等领域展现端到端智能科研体系的潜力^[45-48]。

近年来,大数据与人工智能的应用也推动了冰冻圈研究向跨圈层、多尺度耦合及全球协同方向发展,显著提升了对其演变过程及气候反馈机制的理解(图1)^[30-32]。深度学习与机器学习方法能够高效处理大规模遥感数据,实现冰川、冰盖、海冰及冰山的自动识别与动态追踪,其效率较人工操作提高数千倍,为冰冻圈的实时监测提供关

键支撑^[49]。基于AI的模型可用于冰川退缩趋势预测、冰盖稳定性评估及冰川地貌灾害风险量化,为极地科研和应急管理提供科学依据^[50]。依托高分辨率卫星数据构建的地理空间AI基础模型(如NASA与IBM联合开发的模型)进一步增强了长期监测与大规模数据分析能力^[51]。此外,AI与气候模型及深度观测数据的结合,有助于系统评估极地冰盖的动态演化规律与未来趋势,为全球气候变化研究提供重要支撑^[52]。

尽管AI在冰冻圈研究中展现巨大潜力,但仍面临算法偏差、模型黑箱及对高质量数据高度依赖等挑战,可能影响研究结果的可靠性与公平性^[53]。展望未来,随着“AI for Science”理念的持续深化,冰冻圈研究有望迈向智能化与协同化的新阶段,从而进一步推动跨学科融合与创新。

2 冰冻圈大数据研究范式探索

2.1 冰冻圈研究面临的问题

目前,冰冻圈研究的核心方向聚焦于通过耦合人工智能构建复杂系统模型,以此提升冰冻圈关键参数监测能力、优化冰冻圈过程与风险的模拟预测、揭示冻土与气候的反馈机制,并推动“冰-洋-大气”等多圈层耦合研究的深化。但从“认知冰冻圈变化规律、制定气候变化下致利致害适应性对策”的实际需求来看,现有研究仍存在三方面显著不足^[3,4,54]:第一,缺乏跨尺度的冰冻圈整体研究框架,尚未实现“气候强迫-能量与水量收支-冰冻圈演变-灾害风险”的闭环耦合与定量归因,难以系统解释冰冻圈变化的连锁效应与内在驱动机制;第二,冰冻圈灾害链模拟能力不足,对极端事件驱动下“冰川-冰湖-坡面-河谷”多灾种链式演化过程的认知仍较为有限,无法精准捕捉灾害传导的关键节点;第三,冰冻圈风险评估方法偏重静态分区,对评估过程中的不确定性量化不足,导致评估结果难以支撑动态风险防控决策。

此外,从冰冻圈模型的发展与应用现状来

看,现有物理模型大多局限于区域尺度,难以实现跨区域、多要素的联动模拟;即便全球气候模型包含部分冰冻圈过程,其参数简化处理也导致对冰冻圈反馈过程的刻画不够充分^[26,30]。更为关键的是,目前冰冻圈大模型的构建仍处于空白状态,物理模型与AI方法的耦合能力有限,难以满足未来冰冻圈风险预测与防灾减灾的实际需求^[30]。不过,近年来大数据驱动与人工智能技术的快速发展,已为构建“AI+物理模型”的新型研究范式提供了可行路径。

2.2 冰冻圈“大数据+AI+模型”耦合研究范式

冰冻圈作为气候系统的关键敏感组分,其多要素非线性联动、跨圈层耦合及长灾害链演化过程对全球变化响应机制与风险预测构成重大挑战。传统物理模型在描述复杂过程时虽具备物理一致性,但在应对多源异构数据、参数不确定性和高维非线性动力学时存在明显局限。因此,亟需探索融合“大数据(Data)+人工智能(AI)+模型(Model)”的新型研究范式(DAM),突破现有模拟框架的瓶颈。

在DAM范式中,主要存在三类典型的耦合方式。第一种是“物理模型嵌入AI”路径,其核心在于将守恒律、能量平衡及动力学方程等物理约束引入深度神经网络,从而规避“黑箱”学习的局限,显著提升人工智能在外推预测中的物理合理性与可解释性。第二种是“AI嵌入物理模型”路径,通过利用深度学习技术替代或增强传统模型中计算开销巨大的参数化过程(如积雪-冰川能量收支、冻土热传导等),能有效提高大尺度模拟的效率和精度。第三种是“数据驱动耦合AI的参数优化”路径,其依托多源观测数据并结合人工智能的模式识别与优化能力,动态校准物理模型的关键参数与边界条件,以此缓解参数不确定性对模拟结果的制约。三条耦合路径并非相互孤立,它们可在不同应用层面互为补充,共同为冰冻圈系统的多尺度、多要素一体化模拟提供有力支撑。

基于上述DAM范式,本研究构建了面向全球的冰冻圈研究引擎(Global Cryosphere Engine, GCE),作为耦合范式的工程化落地与集成化实现。该平台整合多源数据处理、智能算法调度与多尺度模拟能力,能够在局地到全球尺度上支撑冰冻圈长期演变的分析与预测。通过引入智能融合计算与自然语言交互机制,GCE为冰冻圈研究提供了高效、灵活且智能化的支撑环境,推动研究范式从单要素模拟向系统化、智能化模拟的跃迁。

3 支持冰冻圈联动研究的GCE引擎研发

3.1 GCE平台基本架构

依托国家重点研发计划项目“冰冻圈大数据挖掘分析关键技术及应用”(2022YFF0711700),团队基于超算与AI智算协同加速的混合云时空计算环境,研发了“冰冻圈研究引擎”(Global Cryosphere Research Engine, GCE)。GCE整合多源全球遥感数据与冰冻圈领域区域特色数据,涵盖冰川、冻土、沙漠、积雪、黄土及灾害监测等数据类型,同步集成AI样本数据集、冰冻圈专业分析算法、物理过程模型及人工智能模型。通过创新的云集成建模框架,GCE支持数据集成、模型调用与模型构建的可视化操作,同时支持计算任务的智能化编排,最终构建起“大数据+AI+模型”的冰冻圈研究新范式及应用场景,GCE基本架构如图2所示。

GCE平台采用分层架构设计,实现冰冻圈研究的全流程智能化支持。底层为算力与存储基础,包括GPU计算资源、超算节点(DCU)及超算框架(MPI、OpenMP),并通过对象存储与云存储提供高性能数据管理。中间层由容器化服务(Kubernetes)和AI框架(PyTorch、TensorFlow)支撑,实现计算资源与任务的弹性调度。AI调度层通过任务池、资源装配与负载管理,结合DAG编排、流程控制与数据流转机制,实现任务的智能分配与执行。上层功能包括模型构建与部署

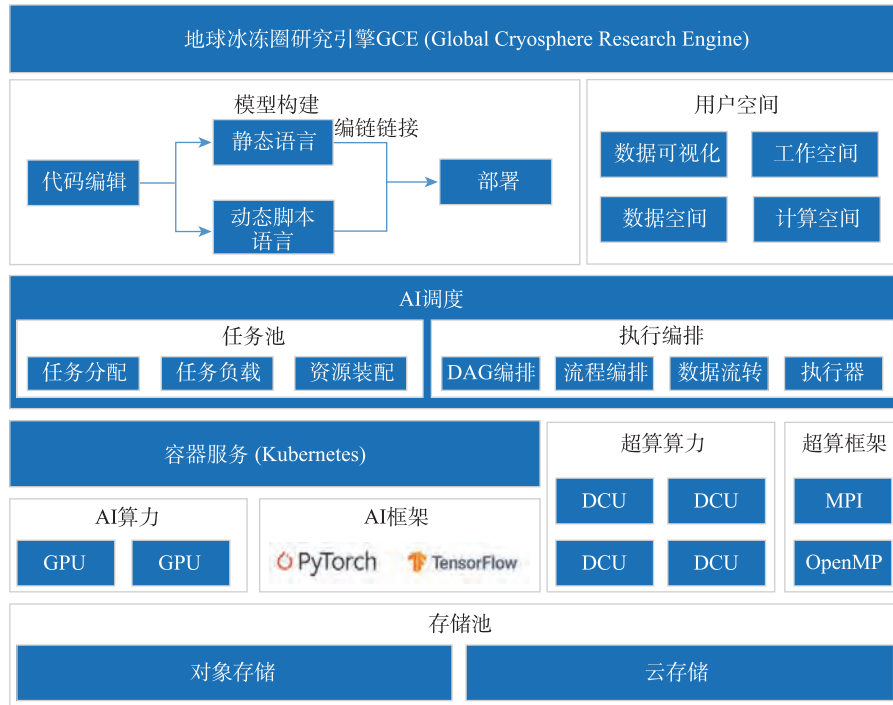


图2 冰冻圈研究引擎GCE的基本架构

Fig.2 Basic Architecture of the GCE

模块,支持静态语言与动态脚本语言的开发与编译,同时提供用户空间,包括数据可视化、工作空间、数据管理及计算环境,实现数据、计算与交互的全面解耦。整体架构体现了从底层算力到顶层智能交互的全链条整合,为多源数据处理、模型模拟与AI辅助科研提供高效、可扩展的技术支撑。

3.2 GCE平台功能与特点

GCE平台以“大数据+AI+模型”的全链条整合为总体目标,通过三大功能集群的深度协同,构建了面向冰冻圈研究的智能化支撑体系(平台界面如图3所示)。这一体系涵盖数据治理、模型计算、智能交互等核心环节,形成从数据到知识的完整科研闭环。

数据治理与制备构成平台基础能力。GCE对1.1 PB冰冻圈数据(全球遥感、观测、特色及AI数据集)实现统一接入、规范化处理、分级管理与安全共享。用户可利用内置工具或自定义脚本

进行数据清洗、重采样和投影转换。平台的专题数据自动制备模块结合算法库与AI调度系统,实现冰川、积雪和冻土等要素的自动化提取与标准化专题数据集生成,确保从多源数据采集到专题产品制备的闭环流程,为后续模型模拟提供高质量输入。

模型构建与流程编排提供核心计算引擎。GCE集成100余种物理及AI融合模型(如OGGM、Noah-MP),支持Fortran、C/C++、Python、Julia等多语言编译与运行,并可进行情景模拟、参数化和批量任务自动执行。模型输出与可视化模块联动,用于误差评估与灵敏度分析,实现模型全生命周期管理。平台内置300余种遥感、水文和GIS算法与数据处理管道深度耦合,形成标准化工具链。流程编排模块基于DAG驱动与AI调度,实现数据预处理、模型运算及AI分析的自动化执行,大幅提升冰川演变预测等任务的效率与稳定性。

AI支持与交互功能强化平台差异化优势。GCE智能体模块深度集成大语言模型(如Deep-

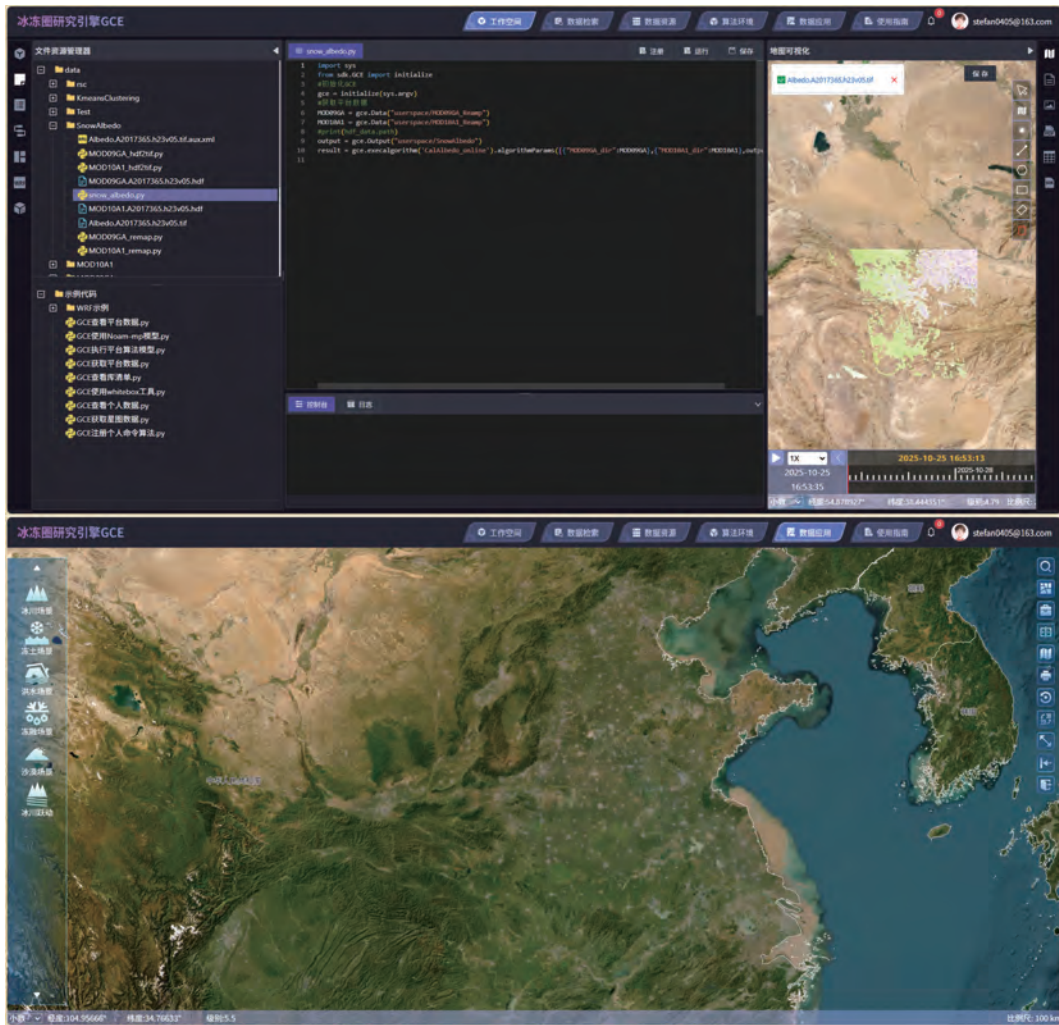


图3 冰冻圈研究引擎GCE的界面

Fig.3 Interface of the GCE

Seek),并嵌入 DAG 流程,为算法推荐、代码生成、参数优化及结果解释提供动态辅助。开发环境支持 GCE-SDK、多语言编辑器及容器化运行,实现计算、数据与工作空间解耦,便于模型调试与快速迭代。

GCE通过异构算力调度、多语言协同架构及多源数据治理等技术的突破,解决了冰冻圈研究中算力割裂、语言壁垒及数据孤岛问题,构建了“技术基座-数据中枢-智能赋能”三位一体的创新生态体系,实现了大规模多源异构数据的全链路治理与标准化处理,支持跨时空尺度的综合联动分析,并在模型选择、参数优化及误差诊断等

关键环节提供智能化辅助,从而推动冰冻圈研究从传统数值模拟向全流程智能模拟的全面转型。

4 “大数据+AI+模型”冰冻圈应用示例

目前,GCE已在国家冰川冻土及沙漠科学数据中心 (<https://www.ncdc.ac.cn>)实现常态化应用,为冰冻圈数据集制备、典型灾害演变分析及关键要素态势感知提供支持。本文围绕青藏高原受气候暖湿化影响,冻土退化迅速显著改变了土壤水热状况这一事实,基于GCE开展了“观测数据+深度学习(分层LSTM)+物理模型

(Noah-MP)”的青藏高原土壤水分融合模型模拟框架构建,系统阐述了“大数据+AI+模型”研究范式在冰冻圈研究过程中的技术流程与实施步骤。

传统陆面过程模型 Noah-MP 瞬态陆面过程模型通过求解非线性温度与水分传导方程模拟冻土水热过程, Noah LSM 在土壤水热的模拟中, 土壤热量模拟较好, 由于参数不确定性及冻土冻融物理机制描述不足导致土壤水分模拟结果不准确。同时, 其传统参数获取方式依赖查找表或传递函数(PTF, Pedotransfer Functions), 存在区域适用性差及空间异质性不足的问题。为提升高原冻土模型预测性能, 本研究提出将长短期记

忆神经网络(LSTM)嵌入 Noah-MP 模型, 通过替代模型中的土壤水分传导方程, 实现物理模型与深度学习方法的深度耦合, 对应于 DAM 范式的第二类耦合路径, 即利用深度学习技术优化传统模型参数化过程, 进而提升大尺度冻土水热过程模拟的效率与精度。

研究全流程在 GCE 中的实现如图 4 所示, 涵盖四个核心环节: (1) 数据集制备: GCE 调用 Noah-MP 生成参数扰动下的高原土壤温湿度场, 并基于 2010—2015 年日尺度中国气象强迫数据集(ITPCAS)及站点观测数据(QT01-QT09)构建多维时序训练样本。训练样本按 4:1 比例划分为训练集与测试集, 用于后续代理模型的训练与验

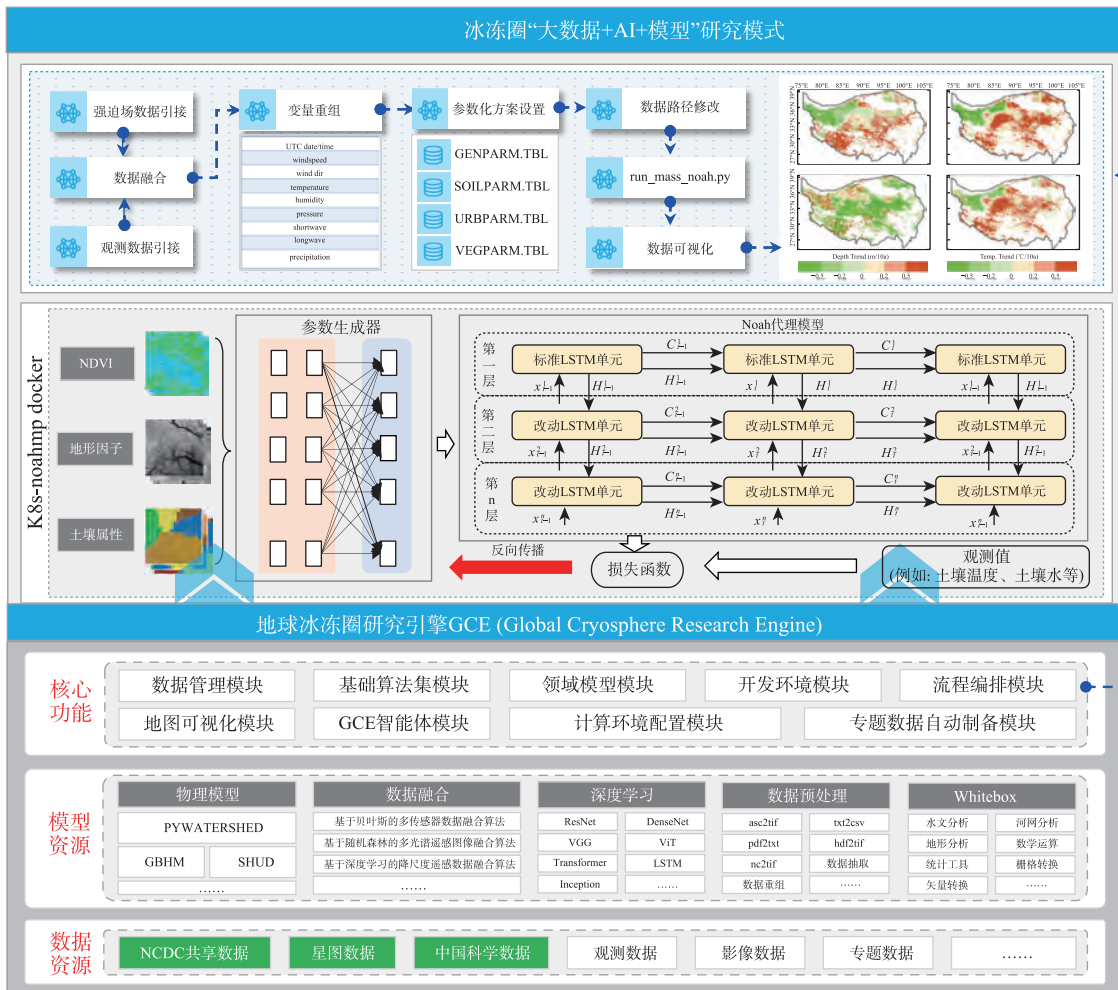


图 4 冰冻圈研究引擎GCE应用示范
Fig.4 Application demonstration of the GCE

证;(2)代理模型构建:在 GCE 分布式 GPU 环境下,改进 LSTM 网络结构并引入注意力机制,构建分层 LSTM 代理模型。该模型绕过复杂物理建模,通过学习土壤水热的非线性关系,精准捕捉冻融作用下的水热耦合机制。分层结构可模拟垂直方向的水分运移,将上一土壤层水分作为下一层输入,有效还原层间水分传输规律;(3)神经网络训练:基于 PyTorch 构建可微计算图,通过训练数据反向传播迭代优化分层 LSTM 参数,捕捉冻融作用下的水热耦合机制,学习土壤水分动态的非线性关系;(4)模型验证:采用皮尔逊相关系数(PCC)、系统偏差(Bias)及均方根误差(RMSE)评价代理模型与原始 Noah-MP 在土壤水分模拟中的性能。在 GCE 平台中,土壤水分模拟对比结果如图 5 所示,表明分层 LSTM 在整体性能上表现更优,其模拟结果与实测数据的相关性平均值超过 0.8,显著高于传统 Noah-MP 模

型。此外,在所有观测站点的模拟中,分层 LSTM 模型的均方根误差(RMSE)均低于传统的 Noah-MP 模型,充分表明该代理模型在模拟精度和泛化能力方面均具有显著优势。

最终,GCE 通过统一编译框架,将上述过程中涉及的 Python 深度学习模块与 Fortran 物理模型协同封装,生成容器化计算单元,并在 GCE 平台上实现流程自动化编排,从数据采集、训练、模拟到验证形成完整闭环,为冰冻圈多要素模拟提供一种可复现、高效且智能驱动的研究范式。

5 结论与展望

针对冰冻圈关键科学问题与应用需求,本文系统梳理了冰冻圈研究范式的发展脉络,重点阐述了冰冻圈在第四范式、人工智能技术加持下的研究范式演变,提出了基于“大数据+AI+模型”

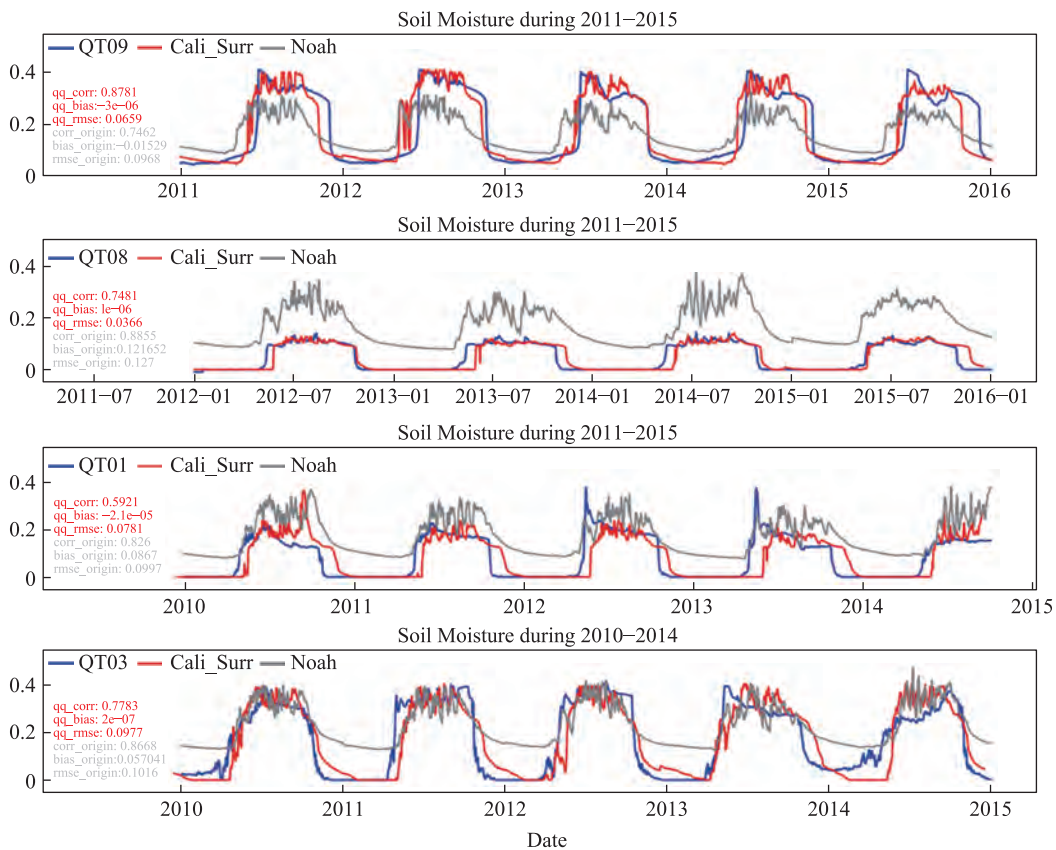


图5 土壤水分模拟精度对比

Fig.5 Comparison of soil moisture simulation accuracy

的研究新范式。同时,依托国家重点研发计划项目,研发了冰冻圈研究引擎(GCE),该引擎整合超级计算与智能计算资源,接入多源遥感及观测数据,提供多语言兼容的一体化编译环境,实现了物理模型与AI方法的高效耦合。基于GCE构建了“观测数据+深度学习+物理模型(Noah-MP)”的青藏高原土壤水分融合模型模拟框架,验证了本文提出的研究范式能支持多要素联动研究,可有效提升冰冻圈模拟精度,也表明GCE平台支撑冰冻圈“大数据+AI+模型”耦合研究范式的可行性与有效性。

目前,该范式虽为冰冻圈多要素联动、跨区域及跨学科的大尺度研究提供了可行路径,但仍面临三方面挑战:一是多源异构数据整合难度大,二是跨尺度数据转换效率不足,三是非结构化数据向“AI-Ready数据集”的转化技术尚处初级阶段。亟需通过集成遥感、地面观测、再分析数据与知识文本,结合物理约束与弱监督学习技术,构建支撑冰冻圈要素识别、过程模拟与风险评估的链式技术体系。未来,通过开发具备自主适应能力的多任务智能体系统,该范式有望构建融合领域知识的多模态冰冻圈大模型,进而实现区域联动与多要素协同的智能监测及快速变化预警,最终成为推动冰冻圈研究范式升级的核心工具。

致谢:

本论文得到了国家冰川冻土沙漠科学数据中心的支持;得到了王丽莉、徐欣钰、艾鸣浩、李红星、赵彦博等人帮助查阅资料以及开展数据工程的实践支持;得到了南京师范大学、三和数码测绘地理信息公司、河南德帆高科软件公司利用数据工程开展业务的应用验证支持。同时,论文也参考引用了部分网络上的论述内容,在此一并表示感谢!

利益冲突声明

所有作者声明不存在利益冲突关系。

参考文献

- [1] 周成林. 冰冻圈科学常见误用专业术语解析[J]. 冰川冻土, 2021, 43(6): 1904-1911.
- [2] 丁永建, 效存德. 冰冻圈变化及其影响研究的主要科学问题概论[J]. 地球科学进展, 2013, 28(10): 10.
- [3] 辛羽飞, 卞林根. 全球冰冻圈变化预测研究现状[J]. 极地研究, 2008, 20(3): 12.
- [4] 丁永建, 杨建平, 方一平, 等. 冰冻圈变化的适应框架与战略体系[J]. 冰川冻土, 2020, 42(1): 11-22.
- [5] 杨建平. “美丽冰冻圈”的缘起与发展[J]. 气候变化研究进展, 2024, 20(6): 711-720.
- [6] 周幼吾, 杜榕桓. 青藏高原冻土初步考察[J]. 科学通报, 1963 (2): 60-63.
- [7] 施雅风. 祁连山冰雪利用研究初步开展[J]. 科学通报, 1958(18):574-575.
- [8] 秦大河, 姚檀栋, 丁永建, 等. 面向可持续发展的冰冻圈科学[J]. 冰川冻土, 2020, 42(1):1-10.
- [9] 张方俭. 我国海冰的基本特征[J]. 海洋科技资料, 1979 (6): 99-125.
- [10] 周尚哲, 赵井东, 王杰, 等. 第四纪冰冻圈——全球变化长尺度研究[J]. 中国科学院院刊, 2020, 35(4): 475-483.
- [11] TUKEY J W. The future of data analysis[J]. *Annals of Mathematical Statistics*, 1962, 33(1): 1-67.
- [12] 李新, 车涛, 李新武. 冰冻圈遥感学[M]. 北京: 科学出版社, 2020.
- [13] 冉有华, 李新, 车涛, 等. 中国冰冻圈遥感近期研究进展与若干前沿问题探讨[J]. 遥感学报, 2025, 29 (6): 1831-1847.
- [14] MUGUNTHAN J S, DUGUAY C R, ZAKHAROVA E. Machine learning based classification of lake ice and open water from Sentinel-3 SAR altimetry waveforms [J]. *Remote Sensing of Environment*, 2023, 299: 113891.
- [15] ZEMP M, HUSS M, THIBERT E, et al. Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016[J]. *Nature*, 2019, 382: 386.
- [16] HUGONNET R, MCNABB R, BERTHIER E, et al. Accelerated global glacier mass loss in the early twenty-first century[J]. *Nature*, 2021, 726: 731.
- [17] 吴小波, 南卓铜, 王维真, 等. 基于Noah陆面过程模型模拟青藏高原植被和土壤特征对多年冻土的影响[J]. 冰川冻土, 2018, 40 (2): 279-287.

- [18] OBU J, WESTERMANN S, BARTSCH A, et al. Northern hemisphere permafrost map based on TTOP modelling for 2000-2016 at 1 km² scale[J]. *Earth-Science Reviews*, 2019, 193: 299-316.
- [19] SUN Z Q, WANG S J, YAN X G, et al. Glacier changes and their impact on glacial lakes in the Parlung Zangbo Basin, Southeastern Qinghai-Tibetan Plateau, 1987-2023[J]. *Journal of Hydrology*, 2025, 661: 133516.
- [20] 王梓霏, 柯长青. 基于深度学习的 Sentinel-1A 影像冰川识别[J]. *遥感信息*, 2022, 37 (4): 43-50.
- [21] 范吉延, 柯长青, 姚国慧, 等. 基于深度学习的全极化 SAR 影像冰川边界识别[J]. *遥感学报*, 2023, 27 (9): 2098-2113.
- [22] HUANG L, LUO J, LIN Z, et al. Using deep learning to map retrogressive thaw slumps in the Beiluhe region (Tibetan Plateau) from CubeSat images[J]. *Remote Sensing of Environment*, 2020, 237: 111534.
- [23] HU J, ZHANG T, ZHOU X, et al. A glacial lake mapping framework in high mountain areas: a case study of the Southeastern Tibetan Plateau[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 1-12.
- [24] 叶世榕, 罗歆琪, 南阳, 等. 一种改进的星载 GNSS-R 卷积神经网络海冰检测方法[J]. *武汉大学学报(信息科学版)*, 2024, 49(1): 90-99.
- [25] 张耀南. 数据工程学建设思考与实践[J]. *数据与计算发展前沿*, 2022, 4 (1): 5-19.
- [26] 王磊, 刘虎, 雍斌, 等. 陆地冰冻圈水文过程的研究现状及展望[J]. *北京师范大学学报(自然科学版)*, 2023, 59(3): 489-496.
- [27] 冷疏影, 丁永建. 自然科学基金资助下的我国冰冻圈科学发展[J]. *地球科学进展*, 2010, 25(10): 1091-1100.
- [28] ABDALLA S, KOLAHCHI A A, ABLAIN M, et al. Altimetry for the future: Building on 25 years of progress[J]. *Advances in Space Research*, 2021, 68: 319-363.
- [29] ZAHOOR A, MAO G, JIA X, et al. Global research progress on mining wastewater treatment: a bibliometric analysis[J]. *Environmental Science: Advances*, 2022, 1: 92-109.
- [30] 王宁练, 刘时银, 吴青柏, 等. 北半球冰冻圈变化及其对气候环境的影响[J]. *中国基础科学*, 2015, 17 (2): 9-14.
- [31] NIU G Y, YANG Z L, MITCHELL K E, et al. The community Noah land surface model with multi-physics options, part 1: Model descriptions and evaluation with local-scale measurements[J]. *Journal of Geophysical Research*, 2011, 116: D12109.
- [32] GOCHIS D J, BARLAGE M, DUGGER A, et al. The WRF-Hydro modeling system technical description[R]. NCAR Technical Note, 2018, 107.
- [33] 邓仲华, 李志芳. 科学研究范式的演化——大数据时代的科学研究第四范式[J]. *情报资料工作*, 2013 (4): 5.
- [34] WAN J B, ZHANG F, PAN J F. Promoting organized basic research: Strategic layout and strategic capacity in science and technology[J]. *Bulletin of Chinese Academy of Sciences (Chinese Version)*, 2021, 36: 1404-1412.
- [35] ZHANG W, LI C, PENG H, et al. CTCNet: A CNN Transformer capsule network for sleep stage classification[J]. *Measurement*, 2024, 226: 114157.
- [36] LIU J, ZHANG Y, LIU J, et al. Automated Recognition of Snow-Covered and Icy Road Surfaces Based on T-Net of Mount Tianshan[J]. *Remote Sensing*, 2024, 16: 3727.
- [37] 金文静. 青藏高原冰川消退区域植被变化格局、过程与驱动因素研究[D]. 长春: 东北师范大学, 2024.
- [38] 宋轩宇. 基于机器学习的典型冰冻圈流域水文过程模拟研究[D]. 兰州: 兰州交通大学, 2023.
- [39] 宋轩宇, 许民, 康世昌, 等. 基于机器学习的冰冻圈典型流域水文过程模拟研究[J]. *地学前缘*, 2023, 30 (4): 451-469.
- [40] 李牧南, 王雯殊. 基于文本挖掘的人工智能科学主题演进研究[J]. *情报杂志*, 2020, 39(6): 82-88.
- [41] 丁璟韬, 徐丰力, 孙浩, 等. 人工智能驱动的复杂系统研究前沿[J]. *电子科技大学学报*, 2024, 53(3): 455-461.
- [42] 王飞跃, 缪青海. 平行科学: 大模型时 AI4 的前沿技术与框架体系[J]. *学术前沿*, 2024(14): 64-79.
- [43] 戴奇乐, 郭金阳, 高阳, 等. 人工智能研究的热点、演进脉络与未来展望——基于文献计量的分析[J]. *价格理论与实践*, 2024 (10): 221-225.
- [44] 常捷. 云计算与人工智能驱动下的数据可视化革新[J]. *中国高新科技*, 2024(17): 25-27.

- [45] 郭蕾蕾. 生成式人工智能驱动教育变革: 机制、风险及应对——以 DeepSeek 为例[J]. 重庆高教研究, 2025, 13(3): 38-47.
- [46] 吕凤先, 刘小平, 陶冶宇. 人工智能驱动的材料化学研究国际发展态势分析[J]. 科学观察, 2025, 20(2): 15-27.
- [47] 吴依洋, 王南男, 熊萍, 等. 人工智能驱动的药物递送: 革新与挑战[J]. 中国科学: 化学, 2025, 55(6): 1623-1634.
- [48] 何斌. 人工智能驱动的网络信息处理与数据分析技术研究[J]. 家电维修, 2025(6): 82-84.
- [49] SHANKAR S. An Artificial Intelligence and Remote Sensing Approach to Iceberg Distribution Around the Greenland Ice Sheet[D]. Lawrence: University of Kansas, 2022.
- [50] UDDIN F, ZAKIR K. Glacial Recession processes and Glacial Lake inventories in High Mountain Asian Communities[C]. XI International Conference on Information Technology and Nanotechnology, Antalya, Turkey, 2025.
- [51] BLUMENFELD J. NASA and IBM openly release geospatial AI foundation model for NASA Earth observation data[EB/OL]. [2023-8-3]. <https://www.earthdata.nasa.gov/news/nasa-ibm-openly-release-geospatial-ai-foundation-model-nasa-earth-observation-data>.
- [52] VASHISHTHA A, MILILLO P, BECKER CAMPOS A, et al. AI-Driven TanDEM-X Penetration Bias Estimation in Antarctica Using ICESat-2 and ECMWF Data: Implications for the NASA Surface Topography and Vegetation Decadal Survey Incubation study [J]. EGUsphere, 2025:1-45.
- [53] 何水兵. 人工智能驱动科研范式变革[J]. 信息化建设, 2025 (4): 17-18.
- [54] 魏亚强, 陈依然, 陈玉玲, 等. 人工智能驱动的地下水数值模拟研究进展[J]. 西北大学学报(自然科学版), 2025, 55(3): 647-657.

收稿日期: 2025年10月10日

张耀南, 中国科学院西北生态环境资源研究院, 研究员, 博士生导师, 国家冰川冻土沙漠科学数据中心主任, 主要研究方向为地学数据工程及数据工程防灾减灾、基于高性能计算环境的地质模型模拟、遥感图像处理及多源数据融合。本文主要承担工作为论文构思、论文撰写。



ZHANG Yaonan, Ph.D., is a professor in Northwest Institute of Eco-Environment and Resources and director of the National Cryosphere Desert Science Data Center. His main research interests include data engineering, disaster prevention and reduction with data engineering, integrated modeling, remote sensing image processing, and multi-source heterogeneous data fusion.

In this paper, he is mainly responsible for the conception and writing of the paper.

E-mail: yaonan@lzb.ac.cn

引文格式: 张耀南, 刘景琦, 康建芳, 南卓铜, 田文彪, 敏玉芳, 赵书萍, 王保得. 冰冻圈“大数据+AI+模型”耦合研究范式探索[J]. 数据与计算发展前沿(中英文), 2026, 8(2): 3-14. DOI:10.11871/jfdc.issn.2096-742X.2026.02.001. <https://cstr.cn/32002.14.jfdc.CN10-1649/TP.2026.02.001>.

ZHANG Yaonan, LIU Jingqi, KANG Jianfang, NAN Zhuotong, TIAN Wenbiao, MIN Yufang, ZHAO Shuping, Wang Baode. Exploration of An Integrated “Big Data+AI+Modeling” Research Paradigm for Cryosphere Studies[J]. Frontiers of Data & Computing, 2026, 8(2): 3-14. DOI:10.11871/jfdc.issn.2096-742X.2026.02.001. <https://cstr.cn/32002.14.jfdc.CN10-1649/TP.2026.02.001>.