科学通报 2024 年 第 69 卷 第 34 期: 5016 ~ 5019
“2023年度中国科学十大进展”解读 观 点

人工智能大模型为精准天气预报带来新突破

张峰[1], 黄小猛[2*], 穆穆[1], 秦博[1], 李佳皓[2]

1. 上海市海陆气界面过程与气候变化重点实验室, 复旦大学大气与海洋科学系/大气科学研究院, 上海 200438;
2. 清华大学地球系统科学系, 地球系统数值模拟教育部重点实验室, 北京 100084
* 联系人, E-mail: hxm@tsinghua.edu.cn

黄小猛 清华大学地球系统科学系长聘教授, 国家超级计算无锡中心总工, 地球系统数值模拟教育部重点实验室主任, 世界天气研究计划中国委员会“大数据和人工智能”工作组组长. 长期从事地球系统高性能计算、大数据和人工智能方面的研究工作, 多项成果被地球系统数值模拟科学大装置和美国大气科学研究中心的CESM模式采用.

天气预报不仅是国际科学前沿领域, 也具有巨大的经济与社会价值. 目前的数值天气预报(numerical weather prediction, NWP)范式可以追溯到20世纪50年代, 通过求解描述大气运动的偏微分方程来推断未来大气状态, 且针对未来几天的预报往往需要花费超级计算机的数百个核时. 欧洲中期天气预报中心(European Centre for Medium-Range Weather Forecasts, ECMWF)的综合预报系统(Integrated Forecasting System, IFS)技巧在众多国家或机构的业务预报模式中脱颖而出, 引领了NWP发展的风向标. 世界气象组织将数值天气预报革命评价为20世纪最重要的科学、技术和社会进步之一.

然而, 使用传统方法进一步提升天气预报水平遇到了瓶颈, 而人工智能(artificial intelligence, AI)技术的不断涌现为解决这些瓶颈提供了新思路[1]. 近年来, 伴随高质量气象资料的不断累积, AI技术, 特别是深度学习方法, 逐渐被用于建模多种多样的地球系统过程, 几乎覆盖了所有时空尺度的天气/气候事件. 同时, AI技术与传统资料同化、集合预报等方法的交叉融合也不断加深. 与传统方法相比, 数据驱动的AI气象模型直接从海量数据中学习潜在物理规律, 而不依赖偏微分方程. 其中, 2023年华为提出的盘古天气AI模型是最具代表性的范例之一[2]. 目前, 大量的回报评估表明, 其准确性可与IFS相媲美, 并且计算成本要小得多, 相关研究论文于2023年7月20日发表在*Nature*. 这种数据驱动建模范式的成功使得AI气象模型成为“2023年中国十大科学进展”中最引人注目的突破之一.

具体来说, 盘古天气定制了3D Earth-specific Transformer架构, 并使用了分辨率为0.25°的40年ERA5再分析资料训练模型, 可以提供包括地势高度、风、压力、温度和湿度等变量的全球三维预报结果. 值得注意的是, 盘古天气较IFS的预报时效约有0.6 d的提升, 且针对台风路径的回报准确性也提高了约25%. 此外, 它可以在几秒钟内生成几天的全球天气预报, 能耗较传统方法降低了10000多倍. 盘古天气也与Google DeepMind的GraphCast[3]和Nvidia的FourCastNet[4]一起被*Science*评选为2023年世界科学十大进展之一, 标志着AI与大数据在天气预报领域取得重大突破.

1 国内外进展

AI气象模型蕴含高度非线性的神经网络结构, 直接从海量数据学习并表征物理过程. 通过数据驱动的方式, 在不依赖明确物理方程的情况下达到甚至超越传统天气预报方法, 引起了“AI+气象”交叉领域的广泛关注, 并掀起了“AI+气象”的建模浪潮. 除了盘古天气之外, 我国涌现出一系列独立发展的AI大气和海洋模型, 如伏羲[5]、风乌[6]、天行[7]、NowCastNet[8]、AI-GOMS[9]和羲和[10]等. 包括复旦大学、同济大学、清华大学、国防科技大学、上海人工智能实验室和国家气象中心在内的许多研究机构也在不断增强这些AI气象模型的预报能力, 例如提升时空分辨率和提供集合预报等. 其中一些模型在一定程度上嵌入了现有的物理规律, 具有较好的物理可解释性, 并且比传统的数值模式具有更高的预报技巧. 国外的许多研究机构也在积极推进AI气象模型的发展, 除了上述GraphCast和FourCastNet外, Google DeepMind的GenCast[11]、Google的MetNet[12,13]系列和NeuralGCM[14]、微软的ClimaX[15]、ECMWF的AIFS均能以纯数据驱动的推理方式超越数值预报技巧.

目前, 包括中国气象局和ECWMF在内的多个全球气象服务机构已经开始实时评估多个AI气象模型的预报性能, 例如盘古天气、伏羲、GraphCast、FourCastNet等, 并致力于将这些模型真正投入业务预报. 这些迹象充分显示出“AI+气

象”前沿研究的强大影响力, 有望以更低的成本为公众提供更高准确率的预报结果. 现有评估结果表明, AI气象大模型具有巨大的潜力, 未来发展值得期待.

2 创新性与科学意义

数据驱动建模范式在提高预报精度和加速推理程等方面展现出明显优势和巨大发展潜力, 并引发起热烈讨论与探索. 例如, AI气象模型是否是大气演变偏微分方程更有效的求解方案? 其预报技巧提升是否源自从海量数据中学到了可靠的物理机制? 如何通过明确的数理方法表征其捕捉到的物理机制? 诸如此类的问题均由于数据驱动建模范式的可解释性不高; 且能够有效回答此类由AI气象大模型衍生出的科学问题和挑战, 可能比AI气象模型本身的价值更大[16].

尽管缺乏明确的可解释性是AI气象模型面临的严峻问题, 但其高技巧和低能耗有望最终服务并受益于公众. 现阶段的AI气象模型可辅助甚至补充传统数值天气预报模型的天气预报信息, 特别是对于传统数值模型难以准确模拟的高影响天气和气候事件.

3 未来探索空间

AI技术正在引领天气预报领域的技术变革. 除了进一步开展“AI+气象”的交叉研究外, 也应注重补强基础研究, 并将AI方法的优势扩展至气象研究的其他领域. 以下是开展进一步探索的几个方面.

(1) 端到端的预报形式. 当前的AI气象模型大多使用ERA5再分析资料训练, 可能不适用于业务系统的初始场, 预报技巧得不到保障. 且大多只支持从固定时刻开始起报, 时效性较差. 可以充分发挥AI模型计算速度快、自带优化模块的优点, 发展基于AI模型的资料同化系统, 发展使用卫星等多源观测资料进行训练的AI气象模型, 可实时获取更新数据并借助AI高效的推理能力提供快速预报, 这对于极端天气的预报和防灾减灾非常有益.

(2) 物理-数据融合驱动的建模. AI气象模型在长时序的预报中可能输出不符合基本物理约束的结果. 可以利用偏微分方程描述明确的动力过程, 在数据驱动的基础上引入物理约束, 使用AI方法建模未知过程, 例如NeuralGCM. 这样的融合建模不仅可以增强AI模型的可解释性, 还有助于提高预报能力.

(3) 耗时过程的AI加速与替代. AI模型的高效性有利于气象领域计算密集型任务的加速, 如集合预报、数据同化和目标观测等; 也可以用来加速数值模式的关键过程, 如参数化方案、辐射传输等. 进一步探索AI方法在改进资料同化、集合预报的计算精度和效率等方面的应用, 有望将获得结果的时间从数小时大大缩短到几分钟.

(4) 引入多圈层变量和过程. 将多圈层(大气圈、水圈、冰冻圈、生物圈、岩石圈)物理要素融入AI模型以开发耦合的AI地球系统模式. 目前, 大多数AI模型只建模单个圈层变量(例如仅大气或海洋), 导致预报时效较短. 然而, 对于较长时间尺度的天气和气候事件, 特别是在次季节到季节尺度, 引入与陆地等其他过程相关的变量是至关重要的. 多圈层的AI模型有望解决当前次季节到季节的预报难题.

(5) 增强泛化能力. AI气象模型的迁移外推成本较低, 然而其训练过程对计算资源的需要庞大. 在计算成本方面, AI气象模型的训练过程与传统数值模式相比并无明显优势. 要提高AI气象模型的应用价值, 需重点关注增强其泛化能力. 此外, 当前AI气象模型面临的过平滑等问题, 使得AI气象模型在预报极端天气事件等方面不尽如意. 根据普遍近似原理[17], 神经网络可精确逼近各种非线性连续算子和泛函. 因此, 研究更具泛化能力的算子学习、生成式模型等方法, 构建面向气象领域的高效AI算法[18,19], 有望提升AI气象模型的泛化性, 进一步拓展AI气象模型的应用空间, 增强对极端事件等的预测技巧.

(6) 基于AI气象模型的可预报性探索. 具有更高模拟和预报能力的AI气象模型必然从海量数据中捕获了可靠的物理机制. 基于这类更高技巧的模型开展天气和气候事件的可预报性研究, 并明确其表征的物理过程, 可进一步提升对地球系统非线性过程的认知和预报技巧, 尤其是对数值模式难以模拟的现象.

(7) 探索AI气象预报与其他新型信息技术融合的新范式. 新兴信息技术如云计算、分布式数据库和数据挖掘等的迅猛发展为AI气象模型的进一步拓展应用提供了更多潜力. 有效融合各种信息技术, 是提升AI气象模型的实际应用水平以及推广其应用的核心. 进一步地, 将AI气象模型与大数据、量子计算、数字孪生等前沿技术相结合, 开拓更广泛的应用领域, 是探索推动AI气象模型应用潜力的有益方向.

4 结语

先进的AI技术可以进一步推动地球科学领域发展, 在一定程度上引起研究范式的转变, 不应简单地将其视为数值模型的替代品. 当前, 数值预报仍遵循传统的发展途径, 从物理原理和基本公式出发, 研究天气气候的变化机制, 并依托数据搜集、加工处理来生成预报结果. 然而, 人工智能驱动的科学研究(AI for science, AI4S)范式的出现[20]为数值预报方法开辟了新的路径. 人工智能算法能够通过海量数据的训练, 透过反向分析输入和输出数据间的因果关系, 揭示数据所包含的规律, 以精确描绘数据背后隐藏的天气气候变化机制. 融合AI气象模型与数值模式的优点, 采用AI气象模型对数值模式的参数化方案等耗时过程进行加速与替代, 在训练AI气象模型的过程中采用数值模式进行物理约束等, 提高预报的可解释性, 都是可以探索的优势互补方法, 从而进一步提高预报技巧. 未来, 应采取更加开放的态度将动力驱动方法和数据驱动方法的优势有机结合起来. 然而, 如何确保AI模型

的可解释性和可信度, 并进一步扩大应用场景以充分发挥其优势, 还需进行更深入的探究.

致谢 感谢国家自然科学基金(42125503, 42288101, 42222506)资助.

推荐阅读文献

1 Huang J B, Chen B. How artificial intelligence is transforming weather forecasting for the future (in Chinese). Chin Sci Bull, 2024, 69: 2336–2343 [黄建平, 陈斌. 人工智能技术在未来改进天气预报中的作用. 科学通报, 2024, 69: 2336–2343]
2 Bi K, Xie L, Zhang H, et al. Accurate medium-range global weather forecasting with 3D neural networks. Nature, 2023, 619: 533–538
3 Lam R, Sanchez-Gonzalez A, Willson M, et al. Learning skillful medium-range global weather forecasting. Science, 2023, 382: 1416–1421
4 Kurth T, Subramanian S, Harrington P, et al. Fourcastnet: Accelerating global high-resolution weather forecasting using adaptive fourier neural operators. In: Proceedings of the Platform for Advanced Scientific Computing Conference, 2023. 1–11
5 Chen L, Zhong X, Zhang F, et al. FuXi: A cascade machine learning forecasting system for 15-day global weather forecast. npj Clim Atmos Sci, 2023, 6: 190
6 Chen K, Han T, Gong J, et al. FengWu: Pushing the skillful global medium-range weather forecast beyond 10 days lead. arXiv: 2304.02948, 2023
7 Mu B, Wang G, Yuan S, et al. TianXing: A linear complexity transformer model with explicit attention decay for global weather forecasting. Adv Atmos Sci, 2024, doi: 10.1007/s00376-024-3313-9
8 Zhang Y, Long M, Chen K, et al. Skilful nowcasting of extreme precipitation with NowcastNet. Nature, 2023, 619: 526–532
9 Xiong W, Xiang Y, Wu H, et al. AI-GOMS: Large AI-driven global ocean modeling system. arXiv: 2308.03152, 2023
10 Wang X, Wang R, Hu N, et al. XiHe: A data-driven model for global ocean eddy-resolving forecasting. arXiv: 2402.02995, 2024
11 Price I, Sanchez-Gonzalez A, Alet F, et al. Gencast: Diffusion-based ensemble forecasting for medium-range weather. arXiv: 2312.15796, 2023
12 Sønderby C K, Espeholt L, Heek J, et al. Metnet: A neural weather model for precipitation forecasting. arXiv: 2003.12140, 2020
13 Espeholt L, Agrawal S, Sønderby C, et al. Deep learning for twelve hour precipitation forecasts. Nat Commun, 2022, 13: 1
14 Kochkov D, Yuval J, Langmore I, et al. Neural general circulation models for weather and climate. Nature, 2024, 632: 1060–1066
15 Nguyen T, Brandstetter J, Kapoor A, et al. Climax: A foundation model for weather and climate. arXiv: 2301.10343, 2023
16 Ebert-Uphoff I, Hilburn K. The outlook for AI weather prediction. Nature, 2023, 619: 473–474
17 Otto S E, Rowley C W. Koopman operators for estimation and control of dynamical systems. Annu Rev Control Robot Auton Syst, 2021, 4: 59–87
18 Mezić I. Analysis of fluid flows via spectral properties of the koopman operator. Annu Rev Fluid Mech, 2013, 45: 357–378
19 Li Z, Kovachki N, Azizzadenesheli K, et al. Fourier neural operator for parametric partial differential equations. arXiv: 2010.08895, 2020
20 Wang F Y, Ding W W. Decentralized science (DeSci): A new paradigm for diverse and sustainable scientific development (in Chinese). Bull Chin Acad Sci, 2023, 38: 1501–1509 [王飞跃, 丁文文. 分布式自主科学: 科学多样可持续性发展的新范式. 中国科学院院刊, 2023, 38: 1501–1509]

Summary for “人工智能大模型为精准天气预报带来新突破”

Large AI models: Pioneering innovations in accurate weather forecasting

Feng Zhang[1], Xiaomeng Huang[2*], Mu Mu[1], Bo Qin[1] & Jiahao Li[2]

[1] *Shanghai Key Laboratory of Ocean-Land-Atmosphere Boundary Dynamics and Climate Change, Department of Atmospheric and Oceanic Sciences/ Institute of Atmospheric Sciences, Fudan University, Shanghai 200438, China;*
[2] *Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China*
* Corresponding author, E-mail: hxm@tsinghua.edu.cn

Weather forecasting not only stands at the forefront of international scientific inquiry but also holds significant economic and societal value. The paradigm of numerical weather prediction (NWP), dating back to the 1950s, relies on solving partial differential equations that describe atmospheric motions to predict future atmospheric states, often requiring hundreds of core hours on supercomputers. The European Centre for Medium-Range Weather Forecasts (ECMWF) and its Integrated Forecasting System (IFS) have notably excelled among various national and institutional operational forecast models, guiding the direction of NWP developments. The World Meteorological Organization heralds the NWP revolution as one of the most important scientific, technological, and social advancements of the 20th century.

However, traditional methods for enhancing forecasting accuracy have reached a plateau, while the emergence of AI technologies offers new avenues to overcome these limitations. In recent years, with the accumulation of high-quality meteorological data, artificial intelligence (AI), particularly deep learning techniques, has been increasingly employed to model diverse Earth system processes across all spatiotemporal scales. The intersection of AI with traditional data assimilation and ensemble forecasting methods is deepening. Unlike traditional approaches, data-driven AI meteorological models learn underlying physical laws directly from extensive data sets without relying on partial differential equations. A prime example of this in 2023 is Huawei’s PanGu Weather AI model, which has demonstrated accuracy comparable to the IFS at significantly lower computational costs. This model’s success was highlighted in a research paper published in Nature on July 20, 2023. The achievements of data-driven modeling have made AI meteorological models one of the most notable breakthroughs in the “2023 Top Ten Scientific Advances in China”.

Artificial Intelligence (AI) meteorological models have garnered significant attention in the “AI + Meteorology” interdisciplinary field by leveraging highly nonlinear neural network architectures to directly learn and represent physical processes from vast data sets. These data-driven models achieve or even surpass traditional weather forecasting methods without relying on explicit physical equations, thus sparking a wave of “AI + Meteorology” modeling initiatives. Apart from the notable PanGu Weather, several independently developed AI atmospheric and ocean models have emerged in China, including FuXi, FengWu, TianXing, NowCastNet, AI-GOMS, and XiHe. Institutions such as Fudan University, Tongji University, Tsinghua University, National University of Defense Technology, Shanghai AI Lab, and the National Meteorological Center are enhancing these AI models’ forecasting capabilities, for instance, by improving spatio-temporal resolutions and providing ensemble forecasts. Some models incorporate existing physical laws to a degree, offering better physical interpretability and forecasting skills than traditional numerical models. Internationally, research institutions are advancing AI meteorological model development, with projects like Google DeepMind’s GenCast, Google’s MetNet series and NeuralGCM, Microsoft’s ClimaX, and ECMWF’s AIFS. These efforts are predominantly data-driven and exceed the numerical forecasting techniques in precision. Global meteorological agencies including the China Meteorological Administration and ECMWF have commenced real-time performance evaluations of multiple AI models like PanGu Weather, FuXi, GraphCast, and FourCastNet, aiming to integrate these models into operational forecasting. These developments vividly demonstrate the profound impact of cutting-edge “AI + Meteorology” research and hold the promise of delivering more accurate forecasts at lower costs. Current evaluations indicate substantial potential in AI-driven large-scale meteorological models, warranting anticipation for their future advancements.

This paper reviews the current research progress on large-scale weather models both domestically and internationally, discusses their scientific significance, and explores future prospects for these models.

artificial intelligence, meteorological big model, weather forecasting, intelligent mode

doi: 10.1360/TB-2024-0543