引文格式：秦昆, 王宇帆, 黄静, 等. 数据驱动与理论驱动双向耦合的智能化地球系统模式探索[J]. 地理空间信息, 2023, 21(3): 1-4. doi:10.3969/j.issn. 1672-4623.2023.03.001

数据驱动与理论驱动双向耦合的智能化地球系统模式探索

秦　昆[1]，王宇帆[1*]，黄　静[2]，刘　娟[2]，汪子正[1]，张　凯[1]，高谢庆[1]
（1. 武汉大学 遥感信息工程学院，湖北 武汉 430079；2. 北京应用气象研究所，北京 100029）

摘　要：分析总结了人工智能方法在地球系统模式中的应用，提出了人工智能数据驱动与物理模型理论驱动相融合的智能化地球系统模式研究框架，并对该框架的研究思路和关键技术进行了分析。该框架将地球系统模式自下而上的数据驱动模型与自上而下的物理模型耦合，通过双向驱动共同实现地球系统数学物理方程的有效求解，从而构建智能化的地球系统模式。进一步研究需要基于该理论框架提出具体可行的算法，从而促进智能化地球系统模式研究的发展。

关键词：人工智能；深度学习；物理模型；人工智能与物理模型耦合；智能化地球系统模式
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2023）03-0001-04

Exploration on Intelligent Earth System Model Coupled Data-driven Models with Theory-driven Models

QIN Kun[1], WANG Yufan[1], HUANG Jing[2], LIU Juan[2], WANG Zizheng[1], ZHANG Kai[1], GAO Xieqing[1]
(1. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China;
2. Beijing Institute of Applied Meteorology, Beijing 100029, China)

Abstract: In this paper, we analyzed and summarized the applications of artificial intelligence in the earth system model, proposed a research framework of intelligent earth system model coupled data-driven models based on artificial intelligence with theory-driven models based on physical models, and analyzed the research thoughts and the related key technologies. Through the double-direction-driven, to realize the effective resolution of mathematical physical equations of earth system, and to build intelligent earth system model. In the future, need to propose concrete and feasible algorithms, and promote the development of intelligent earth system models.

Key words: artificial intelligence, deep learning, physical model, coupled artificial intelligence with physical model, intelligent earth system model

地球系统包含大气圈、水圈、陆地岩石圈、冰雪圈和生物圈之间的复杂相互作用和反馈，涵盖自然和人类的全方位相互作用，涉及数学、地理、物理、化学、生物等学科以及社会和经济等众多领域[1]。地球系统模式是基于地球系统中动力、物理、化学和生物过程建立的数学方程组确定各个部分（大气圈、水圈、冰雪圈、岩石圈、生物圈）的性状，由此构成地球系统的数学物理模型，再采用数值方法进行求解，编制成一种大型综合性计算程序，并通过计算机对地球系统复杂行为和过程进行模拟与预测的科学工具[2-6]。地球系统模式是理解过去气候与环境演变机理、预估未来潜在气候变化情景的工具，世界各发达国家纷纷制定了各种地球系统模式的研究计划，并投入巨资进行研究[2,7]。随着地球系统模式研究的发展，数据量剧增，系统越来越复杂，传统方法已难以满足地球系统模式发展的需求，迫切需要大数据技术和人工智能技术的支撑[1,8]。

将人工智能的机器学习（数据驱动）与物理模型（理论驱动）两种科学范式进行融合，研究地球系统模式的智能化处理和分析技术是一个关键问题。人工智能的深度学习模型借助强大的特征表示和学习能力，日益成为地球系统模式智能化发展的技术基础；

收稿日期：2022-11-30；修回日期：2023-02-21。
项目来源：国家自然科学基金资助项目（42171448）。
第一作者简介：秦昆（1972—），博士，教授，主要研究方向为时空大数据分析、空间人文社会地理计算、地球系统模式，E-mail：qink@whu.edu.cn。
通信作者：王宇帆（2001—），主要研究方向为智能遥感图像处理、地球系统模式，E-mail：EnchantedTay@whu.edu.cn。

但深度学习也存在模型复杂、参数多、学习过程是“黑盒模型”、学习结果可解释性差等问题。因此，将人工智能的深度学习模型（数据驱动）与地球系统模式的物理模型（理论驱动）相融合，将人工智能深度学习自下而上的模型与地球系统模式自上而下的模式紧密耦合，充分利用两种模型的优点，双向驱动实现优势互补，是研制智能化地球系统模式的有效思路之一。本文将系统分析和总结人工智能（机器学习、深度学习等）在地球系统模式中的应用，并进一步提出人工智能（数据驱动）与物理模型（理论驱动）双向驱动的智能化地球系统模式研究框架和方法，最后对智能化地球系统模式的未来发展趋势进行了展望。

1 智能化地球系统模式的方法与应用分析

21世纪以来，人工智能越来越多地应用于地球系统科学，除在固体地球科学领域具有广泛应用以外，在大气和海洋等地球系统模式的其他分量中也有广泛应用。本文重点介绍人工智能方法在陆面过程模式、大气环流和大洋环流3个分量中的应用，并简要分析人工智能在海冰分量、地球化学分量以及地球系统耦合器等方面的应用。

1）人工智能在大气环流分量中的应用。赵宗慈[1]等系统梳理了21世纪以来人工智能（主要是机器学习）应用于气象科学领域的主要国外文献，对人工智能在气象领域的应用进行了总结。人工智能在气象领域的常用算法包括决策树算法、随机森林算法、人工神经网络算法、遗传算法、支持向量机算法、贝叶斯学习算法、Boosting与Bagging算法、关联规则算法以及期望最大化（EM）算法等[1]。众多学者开展了人工智能在大气环流分量中的应用研究，如Coelho等提出了一种利用贝叶斯模型预测厄尔尼诺与南方涛动（ENSO）的方法，该方法有效提升了ENSO预测的准确性和可靠性；SHI X J[9]等将机器学习方法应用于短期临时提前预报，提出了一种卷积长短期记忆（LSTM）模型，即一个端到端的训练模型，该模型能更好地反映时空相关性，取得了很好的降雨量临时预报效果。

2）人工智能在大洋环流分量中的应用。人工智能在地球系统模式大洋环流分量中具有广泛应用。以海浪预测为例，人工智能对海浪预测预报的研究大多是利用机器学习方法（包括神经网络模型），如王燕[10]等利用支持向量回归（SVR）方法建立了渤海海域近岸海浪有效波高短期预测模型，并设计了多组风浪信息组合输入方案，开展了有效波高预测敏感性实验，研究发现，利用SVR方法建立的机器学习模型对小样本数据集具有良好的适应能力，可有效解决预报中的非线性问题，为近岸海浪有效波高短期预测提供参考。在地球系统模式预报中，LSTM网络、循环神经网络、卷积神经网络等较为复杂的深度学习模型主要应用于海浪波高的预报研究，如卢鹏[11]等对传统LSTM神经网络算法进行了改进，提出了一种基于混合模型的海浪高度预测算法，通过太平洋东北海盆海域和马尾藻海域4个站点的浮标数据进行了实验，验证了方法的有效性；莫旭涛[12]等对比分析了循环神经网络、LSTM网络、带卷积的循环神经网络、带卷积的LSTM网络4种网络模型在波高数据建模预测中的效果，结果表明，加入卷积的网络模型的预测结果准确度更高。

3）人工智能在陆面过程模式中的应用。陆面过程是指发生在陆地表层的所有物理、化学、生物过程及其与大气、海洋的相互作用过程。陆面过程模式是指定量描述这些过程以及研究人类活动与环境相互作用的数学物理模式，并可通过计算机实现仿真，是数值天气/气候/地球系统模式的核心组成部分[13]。Pal S[14]等综述分析了机器学习方法在陆面过程模式中的应用，结果表明，利用机器学习方法可提升陆地表面模型的性能，通过改进蒸散发和热通量估计、参数优化等方法可减少不确定性。广泛应用的机器学习方法包括人工神经网络、随机森林模型等[14]。PAN S[15]等对4种基于遥感的模型、两种机器学习模型以及14种陆地表面模型共20种全球陆地蒸散发模型（GTET）进行评估发现，机器学习是GTET的重要方法和未来发展方向，但机器学习方法依赖于训练数据，对于那些缺乏长时间序列数据的区域可能难以发挥有效作用。Everingham Y[16]等将降雨量、温度、辐射和季节性气候指数作为随机森林回归机器学习算法的输入特征，将生物量指数作为输出，对随机森林回归模型进行训练，并通过实验验证了机器学习算法在生物量预测中的重要作用。周佳玮[17]等提出了一种结合高分辨率和多时序遥感影像，基于深度学习的水稻提取方法，即构建全卷积网络对高分辨率遥感影像进行分类，并利用LSTM网络和随机森林分类器对哨兵2号多时序遥感影像进行分类，再通过面向对象的分割和投票对3种方法的分类结果进行融合，得到最终提取结果。

4）人工智能在地球系统模式其他分量中的应用。人工智能方法在地球系统模式的海冰分量、地球化学模式以及耦合器等方面的研究相对较少，如郑冬梅[18]等研究了逐步回归分析方法在渤海海冰等级预报

中的应用，通过相关分析和逐步回归分析发现，渤海海冰的年际变化与西太平洋副高面积和北界指数、太平洋副高面积指数、亚洲区极涡面积指数以及南方涛动指数都有显著的相关关系；尹鹏[19]以深度学习方法为基础，分析了海冰密集度数据集的预测可用性，并基于评估结果开展了LSTM网络模型在海冰面积预测中的研究，提出了适用于多变量海冰面积时间序列预测的深度学习方法。复杂地质条件下地球化学异常的识别与提取是地质大数据和人工智能算法研究应用的热点领域[20-22]。地球化学研究中应用的机器学习算法包括神经网络、支持向量机、随机森林、决策树、极限学习机等。机器学习的各种算法不仅能处理大量与矿产预测相关的证据图层，还具有识别已知矿床与证据图层之间非线性关系的潜力[22-24]。复杂化探数据及其异常经常包含复杂和非线性模式，深度学习对于智能识别与提取复杂地质条件下的地球化学异常具有优异的能力。卷积神经网络是其中较为常用和有效的方法[25-27]。这些深度学习方法将输入的数据映射到少数深层次特征中，有利于保留与复杂地质过程及其相互作用有关的深层次结构表征，最终达到提高异常识别的效果[22]。耦合器是地球系统模式的核心器件之一，用于实现不同分量模式间的相互作用和系统集成，能支持不同分量模式间的数据传递、不同网格之间数据的插值、物理场向“目的分量”的汇合以及能量等的守恒性诊断等[28-30]。目前，还鲜有人工智能在耦合器中的应用研究，未来可以考虑将人工智能方法引入耦合器的研发，基于人工智能的方法自动配置各分量耦合时的物理量和耦合频率等。

2　数据驱动与理论驱动融合的智能化地球系统模式

基于深度学习的数据驱动模型在地球系统模式中得到了很好的应用[31]，如Scher S[32]提出了一种数据驱动的天气和气象预报模型，该模型利用深度神经网络构建一个通用的环流模型，通过利用实际数据对该模型进行训练从而实现有效的气象预报；Simmons A[33]等提出了地球系统科学观测与集成的路线图。基于深度学习的数据驱动模型具有自动分析地球系统模式数据的优点，但存在可解释性差等问题。Reichstein M[31]等指出下一步的发展方向是构建耦合地球系统模式物理过程模型与数据驱动的机器学习模型的智能化地球系统模式。

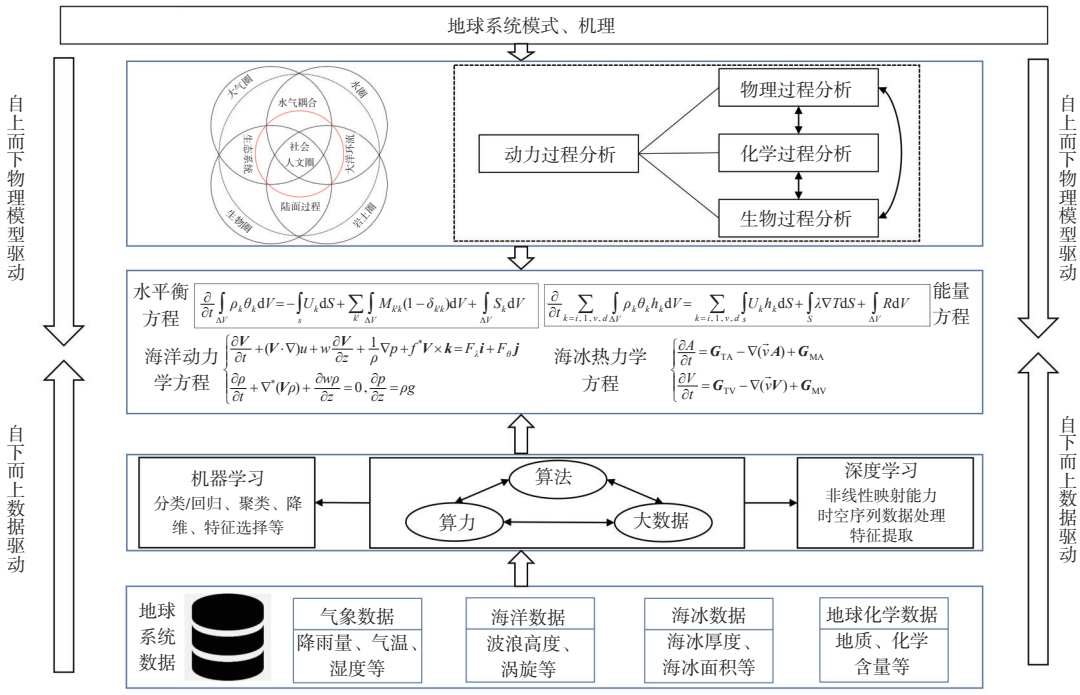


图1　数据驱动与理论模型驱动耦合的智能化地球系统模式框架

针对如何将自下而上基于机器学习的数据驱动模型与地球系统模式自上而下的理论驱动模型进行耦合，从而实现智能化地球系统模式的问题，本文提出了一个理论框架，如图1所示。智能化地球系统模式框架是一种自上而下物理驱动模型与自下而上数据驱动模型耦合的双向驱动模型。

一方面，从理论分析入手，深入研究了地球系统模式的机理模型，地球系统圈层的相互作用，大气圈、水圈、岩石圈、生物圈之间的相互作用及其相互交融形成的特殊圈层——社会人文圈，包括水气耦

合、大洋环流、陆面过程、生态系统等；同时分析了地球系统各圈层相互作用体现出的动力过程、物理过程、化学过程和生物过程等。基于地球系统模式的理论分析，构建地球系统模式的相关数学物理方程，包括水平衡方程、能量方程、海洋动力学方程、海冰热力学方程等。另一方面，从数据分析入手，在收集整理海量地球系统数据（包括气象数据、海洋数据、海冰数据、地球化学数据等）的基础上，利用基于机器学习、深度学习的模型，高性能计算机的算力、精巧的算法以及大数据分析模型等，采用自下而上的数据驱动模型，从海量的地球系统数据中提取信息、挖掘知识，从而求解地球系统模式数理方程中的模型参数。将自上而下的理论驱动模型，与自下而上的数据驱动模型耦合，共同实现地球系统模式数理方程求解任务，从而实现智能化地球系统模式。

3 总结与展望

引入人工智能的理论和方法，发展智能化地球系统模式，是地球系统模式的未来发展趋势。地球系统模式研究中积累了越来越丰富的海量数据，基于机器学习、深度学习等智能化模型，可从地球系统模式大数据中自动探测各种模式、知识和规律，求解地球系统模式的数学物理方程，以及为参数调整等提供支撑。然而，自下而上的数据驱动模型是“黑盒模型”，存在可解释性差等问题。地球系统模式的物理模型是从地球系统模式的机理分析入手，通过分析地球系统圈层相互作用以及地球系统模型的动力过程、物理过程、化学过程、生物过程等，构建地球系统模型的数学物理理论方程。将地球系统模式自下而上的数据驱动模型与地球系统模式自上而下的物理模型进行耦合，双向驱动，共同实现地球系统数学物理方程的有效求解，从而构建智能化地球系统模式，可有效解决这些问题。本文提出了一种耦合双向模型的智能化地球系统模式的理论框架，可为智能化地球系统模式的进一步深入研发提供参考。今后研究需要基于该理论框架，提出具体可行的算法，从而促进智能化地球系统模式研究的发展。

参考文献

[1] 赵宗慈，罗勇，黄建斌，等. 人工智能应用于地球系统科学[J]. 气候变化研究进展，2020，16(1):126-129

[2] 王斌，周天军，俞永强. 地球系统模式发展展望[J]. 气象学报，2008，66(6):857-869

[3] Weber S L. The Utility of Earth System Models of Intermediate Complexity (EMICs) [J]. Wiley Interdisciplinary Reviews: Climate Change，2010，1(2):243-252

[4] Rasch P J，XIE S，MA P L，et al. An Overview of the Atmospheric Component of the Energy Exascale Earth System Model[J]. Journal of Advances in Modeling Earth Systems，2019，11(8):2 377-2 411

[5] Maher P，Gerber E P，Medeiros B，et al. Model Hierarchies for Understanding Atmospheric Circulation[J]. Reviews of Geophysics，2019，57(2):250-280

[6] Danabasoglu G，Lamarque J F，Bacmeister J，et al. The Community Earth System Model Version 2(CESM2) [J]. Journal of Advances in Modeling Earth Systems，2020，12(2):e2019MS001916

[7] 周天军，陈梓明，邹立维，等. 中国地球气候系统模式的发展及其模拟和预估[J]. 气象学报，2020，78(3):332-350

[8] Bergen J B，Johnson P A，De Hoop M V，et al. Machine Learning for Data Driven Discovery in Solid Earth Geoscience[J]. Science，2019，363(6 433):236-245

[9] SHI X J，CHEN Z R，WANG H，et al. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems，2015，1:802-810

[10] 王燕，钟建，张志远. 支持向量回归的机器学习方法在海浪预测中的应用[J]. 海洋预报，2020，37(3):29-34

[11] 卢鹏，年圣全，邹国良，等. 基于变分模态分解和注意力机制的浪高预测[J]. 海洋测绘，2021，41(2):34-39

[12] 莫旭涛，李自立. 神经网络北部湾海洋浪高预测[J]. 现代电子技术，2021，44(13):55-59

[13] 戴永久. 陆面过程模式研发中的问题[J]. 大气科学学报，2020，43(1):33-38

[14] Pal S，Sharma P. A Review of Machine Learning Applications in Land Surface Modeling[J]. Earth，2021(2):174-190

[15] PAN S，PAN N，TIAN H，et al. Evaluation of Global Terrestrial Evapotranspiration by State-of-the-art Approaches in Remote Sensing，Machine Learning，and Land Surface Models[J]. Hydrology and Earth System Sciences，2020，24(3):1 485-1 509

[16] Everingham Y，Sexton J，Skocaj D，et al. Accurate Prediction of Sugarcane Yield Using a Random Forest Algorithm[J]. Agronomy for Sustainable Development，2016，36(2):1-9

[17] 周佳玮，涂理林，陈洪建，等. 融合空间和时序遥感信息的深度学习水稻提取[J]. 地理空间信息，2022，20(2):39-44

（下转第20页）