中国科学: 地球科学 2023 年 第 53 卷 第 3 期: 546 ~ 560
SCIENTIA SINICA Terrae earthcn.scichina.com
《中国科学》杂志社 SCIENCE CHINA PRESS
论 文

地球表层特征参量反演与模拟的机理-学习耦合范式

沈焕锋[1,2†], 张良培[3*]

1. 武汉大学资源与环境科学学院, 武汉 430079;
2. 地理信息系统教育部重点实验室, 武汉 430079;
3. 测绘遥感信息工程国家重点实验室, 武汉 430079
† 通讯作者, E-mail: shenhf@whu.edu.cn
* 通讯作者, E-mail: zlp62@whu.edu.cn

收稿日期: 2022-03-27; 收修改稿日期: 2022-08-09; 接受日期: 2022-09-08; 网络版发表日期: 2023-01-19
国家自然科学基金重点项目(批准号: 42130108)资助

摘要 构建物理驱动的机理模型一直是估算地球表层特征参量的核心科学范式, 发展数据驱动的机器学习模型是地学研究范式转换的重要途径, 而耦合机理模型与学习模型则可以实现“理性主义”与“经验主义”的结合, 是当前最受关注的研究热点之一. 文章针对参量估算的遥感反演与动力学模拟方法, 深入分析了机理模型与学习模型的内在瓶颈及其互补性, 搭建了以机理级联学习、学习嵌入机理、机理融进学习为核心的耦合范式框架, 归纳了预处理与初始化、中间变量传递、后置精化处理、模型替代、模型调整、模型求解、输入变量约束、目标函数约束、模型结构约束、混合应用等十种具体耦合方式, 剖析了当前的主要问题与未来的挑战方向. 研究为深入理解、应用机理-学习耦合模型提供了新视角, 为提升地球表层特征参量反演与模拟能力、服务地球系统科学发展提供理论与技术支撑.

关键词 机理模型, 机器学习, 模型耦合, 遥感反演, 数值模拟

1 引言

地球表层过程的气候变化、环境污染等问题, 深刻影响着人类的生活生产与生命健康. 要深刻理解地球表层复杂的自然与人文现象、促进社会经济的可持续发展, 需要综合、完整和持续的感知数据(中国科学院地学部地球科学发展战略研究组, 2009). 卫星遥感反演与动力学数值模拟是获取宏观、连续地球表层特征参量数据的两个重要手段(陈发虎等, 2019), 如何不断提升遥感反演与数值模拟的精度与能力, 是地球表层系统科学研究的关键基础问题.

无论是遥感反演还是动力学数值模拟, 构建物理可解释的机理模型一直核心的科学范式(De Bézenac 等, 2019). 在遥感反演方面, 基于辐射传输物理过程的定量反演是获取水、土、气、生等多个圈层特征参量的主要途径, 研究学者已发展了大量具有严格物理

中文引用格式: 沈焕锋, 张良培. 2023. 地球表层特征参量反演与模拟的机理-学习耦合范式. 中国科学: 地球科学, 53(3): 546–560, doi: 10.1360/SSTe-2022-0089
英文引用格式: Shen H, Zhang L. 2023. Mechanism-learning coupling paradigms for parameter inversion and simulation in earth surface systems. Science China Earth Sciences, 66(3): 568–582, https://doi.org/10.1007/s11430-022-9999-9

机制的遥感反演方法(李小文, 2005; 梁顺林等, 2016; 李召良等, 2016), 并发布了系列全球及区域尺度的定量遥感参量产品(张正等, 2016). 在数值模拟方面, 各国科学家构建了多种大气数值模型(Skamarock等, 2005)、陆面过程模型(孟春雷和戴永久, 2013)、水文模型(Arnold等, 1998)等, 并基于超级计算平台开发了地球系统模拟器(陈春等, 2005; 邱晨辉, 2021). 总之, 以物理驱动为基础的机理模型是地球表层特征参量反演与模拟的"主框架"(De Bézenac等, 2019), 是地学知识发现的重要基石(Karpatne等, 2017b).

近年来, 地球科学正在经历从数据匮乏到数据丰富的转变(Karpatne等, 2019), 人们获取与生产时空数据的能力已经远大于对其进行处理、分析与理解的能力(Reichstein等, 2019). 在此背景下, 基于大数据的第四科学范式悄然崛起, 并成为地学研究的重要支撑(郭华东等, 2014; 宋长青, 2016; 程昌秀等, 2018; 邓敏等, 2020; 周成虎等, 2021; 李新等, 2022). 以机器学习为代表的人工智能技术发展迅猛, 被认为是挖掘大数据潜力的"金钥匙"(郭仁忠等, 2020; 陈军等, 2021), 在卫星遥感与数值模拟领域受到广泛关注和快速发展(Hsieh和Tang, 1998; 黎夏和叶嘉安, 2005; 宫鹏, 2009; Härter和de Campos Velho, 2010; 张兵, 2018). 在IEEE地学与遥感学会组织的融合分析大赛中, 深度学习模型获得了近年多数赛道的冠军(黄昕等, 2021); 在定量应用方面, 机器学习模型已被广泛地应用于数十种特征参量的遥感反演(郭庆华等, 2020; Yuan等, 2020; 胡斯勒图等, 2020; 冉有华等, 2021). 与此同时, 机器学习也已被成功应用于大气(Navares和Aznarte, 2020)、水文(Petty和Dhingra, 2018)、海洋(De Bézenac等, 2019)等地表过程的模拟预测, 并表现较大的应用潜力. 有鉴于此, 机器学习被期待成为释放数据驱动潜能、加速科学发现的重要框架(Karpatne等, 2017b), 有学者认为它将地学研究推向即将取得重大进展的门槛(Bergen等, 2019).

显而易见, 以大数据为支撑的机器学习模型已经对正统的机理模型形成冲击(裴韬等, 2019), 甚至有学者认为可能导致"理论的终结(the end of theory)"(Anderson, 2008). 然而, 一些学者却坚持当前普遍存在"大数据傲慢"问题(Lazer等, 2014), 机器学习的效能被高估. 例如, 谷歌发布神经网络降水预报模型MetNet (Sønderby等, 2020), 声称在8h的预测中神经网络模型已经优于机理模型, 但是其在学术界受到不少质疑, 至少在长期预测、大尺度预测等方面其仍然不能替代机理模型(Witt等, 2021; Chantry等, 2021) . 针对机器学习的地学应用问题, *Nature*、*Science*等期刊近期相继发表论文(Bergen等, 2019; Reichstein等, 2019; Bauer等, 2021), 认为地学过程的复杂性、交互性、多尺度特性, 以及数据的不确定性、真实样本的稀缺性等, 使得机器学习模型仍然不能替代机理模型, 但是两种模型具有天然的互补优势, 耦合机理模型与学习模型是极具前景的发展方向.

然而, 将显式的机理模型与隐式的学习模型进行耦合存在诸多挑战, 尽管当前已经取得了一些研究进展, 但仍然缺乏标准、统一的范式框架, 导致相关研究模式各异、缺乏关联, 甚至出现相互混淆的问题. 本文在充分归纳现有工作基础之上, 力图构建系统的机理-学习耦合范式框架, 剖析不同耦合方式的特点与潜力, 并展望未来的挑战方向, 以期为相关研究提供理论与应用参考, 促进地球表层特征参量遥感反演与数值模拟技术发展, 为提升地球表层特征参量估算能力、服务地球系统科学发展提供理论与技术支撑.

2　两种模型的优势与瓶颈

2.1　机理模型

"机理"可以广义地理解为任何表达地理对象属性或要素之间有效关系的知识(von Rueden等, 2023), 既包括物理知识, 也包括几何约束、地学规律等. 机理模型遵循客观规律建立输入与输出之间的显式关联, 帮助人们认识与理解所生存的物理世界(Karpatne等, 2017b). 经典的定量遥感反演方法基于大气辐射传输等模型, 将对地观测的电磁波信号与特征参量建立关联, 实现对地球表层的面域感知; 数值模拟系统通过其内在物理过程和动力学机制, 获得地理对象在时间和空间上的连续演进(李新等, 2007). 可见, 机理模型可以较为清晰地描述系统的内部特性, 理论严谨、模型(相对)稳定、结果可解释是其突出的优点. 然而, 机理模型也存在其难以克服的不足.

(1) 机理认知局限. 地球表层系统是多要素混杂、多尺度耦合、多过程交织的复杂巨系统(陈旻等, 2021), 现有的机理模型仍然难以实现对所有地学过程的精准刻画, 一些物理过程仍然未知. 例如, 当前仍然缺乏针对很多参数(如气温、$PM_{2.5}$等)的遥感机理反演

模型，数值模型中也并非所有子过程都可以进行精确物理建模, 约简或近似处理往往导致不确定性.

(2) 欠定系统问题. 即使一些地学过程的机理较为清晰, 但参量反演往往是一个欠定系统, 即观测方程个数少于未知数个数, 导致模型求解十分困难, 对此通常需要一些假设条件，而当假设条件与真实不符时就会带来较大的求解误差. 例如，地表温度遥感反演就是利用N个观测值(波段数)解决N+1个未知数(N个地表发射率和地表温度)的病态问题.

(3) 计算负担问题. 一些机理过程的计算量巨大, 例如，在美国大气研究中心的大气模式中，物理过程的计算占到整体模型计算量的70%(Krasnopolsky等, 2005). 如果进一步提升在分辨率、一致性等方面的要求, 计算量又将呈指数上升, 带来较大的应用困扰.

2.2 学习模型

机器学习模型通过“训练”与“预测”模拟人类的“归纳”与“推测”过程，实现对典型问题的建模与求解. 与机理模型的显式表达不同，学习模型通过对数据的训练建立不同变量之间的隐式关联，即往往是典型的“黑箱”模型. 学习模型的关键优势之一就是当机理未知时，可以跳过对物理过程的理解而直接进行数据驱动的建模，特别是在训练数据充足的条件下，往往可以获得较高的建模精度. 此外，机器学习虽然在训练阶段比较耗时，但在测试应用阶段一般具备较高的计算效率, 也成为其重要的优势之一. 尽管如此, 机器学习模型仍然具有诸多局限，特别是在过程复杂的地学应用中经常存在如下问题:

(1) 泛化性不足. 缺乏足够的训练样本是机器学习地学应用中最为常见的问题，而利用有限的样本去学习复杂的地学过程，极易出现过拟合现象，即使在训练样本上表现出较高的建模精度，测试应用精度也会大幅降低. 特别是当实际的数值范围、变量关系等没有被训练样本所涵盖时，预测结果更可能出现极大偏差, 即典型的泛化能力不足问题.

(2) 迁移性不足. 区域性是地理学的本质特征，不同区域之间不仅表现为不同地理要素的差异，更表现为各要素之间关系的差异. 因此，在某一区域训练的机器学习模型往往难以迁移到其他区域进行应用. 其次，地球表层要素及其相互关系也处于不断变化的过程中，人类活动影响使之变化更为剧烈，如此同一区域不同时间跨度的模型也往往难以通用. 此外，尺度迁移性不足也是地学应用中的又一困境.

(3) 可解释性不足. 科学研究的目标不仅在于发展一个可用的模型，更加在于发现不同变量之间的内在因果关系与驱动模式，并用之实现对理论与假设的解释，从而促进科学知识的进步(Karpatne等, 2017a). 机器学习的一个突出问题就是在可解释性方面存在不足, 虽然在特定条件下也可以获得比较高的精度, 但缺乏对内在机理过程的解释能力.

通过以上分析可见，机理模型与学习模型虽然各有其建模优势, 但也都存在难以克服的不足. 显然，二者之间具有天然的互补性(Ganguly等, 2014; 吴志峰等, 2015)，耦合机理模型与学习模型可以实现“理性主义”与“经验主义”的结合，可以有效调整机理模型的“偏见”, 避免学习模型的“傲慢”(Chantry等, 2021), 因此是必然的发展趋势.

3 机理模型与学习模型的耦合范式

机理模型与学习模型的耦合近期成为各领域的研究热点，但实际上从20世纪末开始，无论在数值模拟(Chevallier等, 1999)还是遥感反演(Aires等, 2001)领域，就已有机理模型与学习模型耦合的思想与成功案例，但受认知水平与技术条件的限制，该方向研究并没有得到足够的关注与发展. 直至最近，随着神经网络特别是深度学习技术的再度崛起, 机理-学习耦合已成为包括地学在内诸多领域的研究热点.

近年来，在英文文献中出现了诸多表达机理模型与学习模型耦合的名词术语，可从图1所示的三列之中各任选一词连接起来, 如“*Physics Informed Machine Learning*”等. 然而，以上各种组合表达过于强调“学习”，而把“机理”放在了次要位置. 但实际上二者的耦合模式多种多样，“机理”与“学习”所占比重也各不相同, 最好能够保持二者之间的平衡, 为此Shen等(2022)用“*Coupling of Mechanism and Learning*”进行表达.

本文提出在地学参量反演与模拟中，可将机理模型与学习模型的耦合归纳为三类基本范式: 机理级联学习、学习嵌入机理、机理融进学习(后分别简称为级联、嵌入、融进), 如图2所示. 机理级联学习就是将两种模型进行前后串联，一种模型的输出作为另一种模型的输入. 学习嵌入机理，就是以机理模型为主、

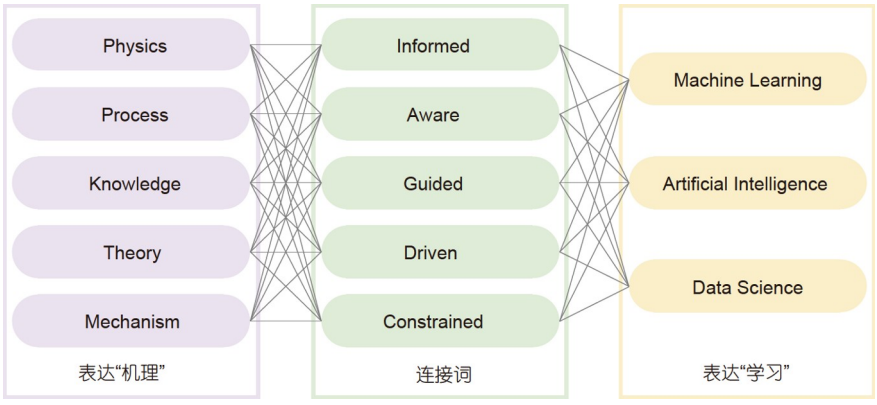

图 1　机理-学习耦合的英文术语

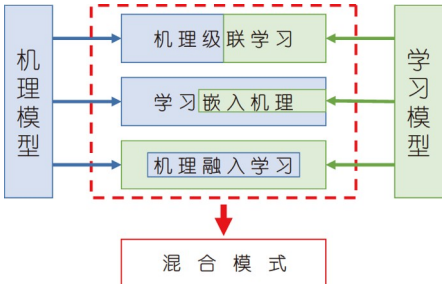

图 2　机理-学习耦合的基本范式

学习模型为辅, 将学习模型嵌入到机理模型之中, 对原有的不确定过程进行替代或优化. 机理融进学习, 就是以学习模型为主框架, 将物理知识融入其中, 从而实现对学习过程的约束引导. 除此之外, 为了发挥三种范式的各自优势, 还可以将它们联合起来, 构建混合应用模式.

3.1　机理-学习级联耦合范式

机理模型与学习模型最为简单的耦合即为级联模式, 通过前后串联、顺序建模的方式实现直接结合. 根据两种模型在整个系统中的应用阶段及其重要程度, 可细为预处理与初始化、中间变量传递、后置精化处理三种具体方式.

3.1.1　预处理与初始化

(1) 质量控制. 利用学习模型对机理模型的输入数据进行质量控制, 可以有效提升后续参量估算的精度. 例如, 遥感数据经常存在噪声、像元缺失等问题, 在基于机理模型的参量反演之前, 首先利用机器学习进行噪声去除、像素补全等处理, 可有效提升机理模型输出的精度与可靠性.

(2) 参数优化. 机理模型的运行往往需要较多输入参数, 模型精度很大程度上受限于输入参数的准确性(张添等, 2012). 机器学习可以被用于获取更加准确的模型参数, 为后续的模型计算提供更优的初始条件. 例如, Beck等(2016)构建了基于回归的水文模型参数局地化方法, 在全球尺度上进行了成功应用; Sawada (2020)利用高斯过程回归模型对陆面过程模型进行了参数优化, 有效提升了模型模拟的精度.

(3) 样本生成. 在很多地学应用中, 往往难以获取机器学习模型所需的真实训练数据, 此时则可借助机理模型生成训练样本. 例如, Aires等(2001)首先利用微波遥感辐射传输方程生成训练数据, 再基于机器学习方法反演了大气水汽、地表温度、发射率等参数. 除此之外, 在基于热红外遥感的地表温度反演(Mao等, 2007)、基于光学遥感的叶面积指数反演(Campos-Taberner等, 2016)、总初级生产力反演(Wolanin等, 2019)、植被含水量反演(Trombetti等, 2008)等应用中, 辐射传输方程也被广泛地用于机器学习的样本构建.

(4) 迁移学习. 为了避免真实观测样本不足导致的过拟合问题, 可退而求其次, 首先利用机理模型生成较粗的训练数据进行预训练, 当模型达到较为稳定的状态后再基于少量的高精度真实样本进行精训练(如图3), 这是迁移学习的一种典型应用形式. Jia等(2021)在预测湖泊水温时, 首先利用基于物理过程的通用湖泊模型生成模拟数据, 并用之进行长短期记忆神经网络

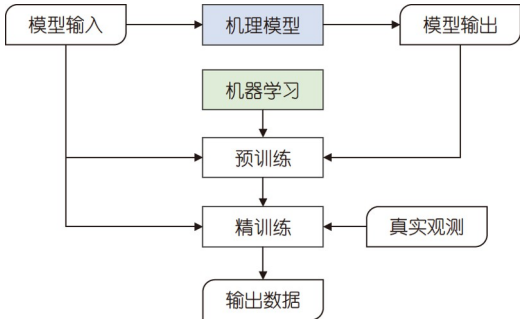

图 3　迁移学习耦合方式

模型的预训练，有效地减少了对真实训练样本的依赖(Read等, 2019).

3.1.2　中间变量传递

受机理认知、技术局限等因素的影响，一些特征参量难以通过完全物理过程的方法获取，此时可通过中间变量的传递实现物理模型与学习模型的联合应用, 即首先基于物理模型估算中间变量, 再利用机器学习模型实现目标参量的估算. 例如, 针对近地气温、大气$PM_{2.5}$等参量，目前仍然缺乏有效的全机理遥感反演模型，但地表温度、气溶胶光学厚度分别与气温、$PM_{2.5}$有较强的相关性，并且已经存在较为成熟的机理反演方法，因此可首先基于机理模型反演地表温度、气溶胶光学厚度，再将反演结果作为机器学习的输入，进而实现气温、$PM_{2.5}$的反演(Shen等, 2018; Shen等, 2020)，如图4所示. 当然，中间变量也可以通过动力学机理模型的数值模拟来获得(Xiao等, 2017)，例如，Liang等(2020)首先基于水质模型模拟六种水质参数，然后将它们输入到长短时记忆神经网络，用于实现叶绿素a含量的预测.

3.1.3　后置精化处理

为了提升遥感反演或动力学模拟等机理模型输出结果的精度、分辨率等指标，可以利用机器学习模型进行后置的精化处理，这也是机理模型与学习模型较为传统的耦合方式之一，具体可包括误差校正、降尺度、集成优化等多种类型.

(1) 误差校正. 基于机器学习的误差校正方法，已广泛应用于遥感反演与模型模拟参量数据的处理，通过建立模型输出数据与地面真实观测或其他参考数据

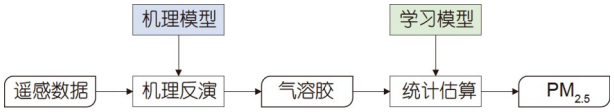

图 4　机理-学习级联PM2.5反演示意图

之间的映射关系，通过后置的校正处理提升原有输出的精度或一致性. Rasp和Lerch(2018)利用神经网络模型进行集合天气预测系统的系统误差校正，无论从精度还是效率方面都比原有模型有了较大提升. Ivatt和Evans(2020)利用梯度提升树模型校正大气化学传输模型的输出，有效提升了臭氧的模拟精度. Noori等(2020)以站点观测为参考，利用机器学习方法对SWAT水文模型的输出进行校正，有效提升了三种关键水质参数的模拟精度.

(2) 降尺度. 大区域尺度遥感反演、模型模拟数据的空间分辨率往往较粗，难以满足精细监测与分析的需求. 在机理模型反演或模拟的基础上，机器学习可被进一步用于降尺度处理，提升数据的空间分辨率. 当前，机器学习已成为遥感反演的降水(Wang等, 2021)、土壤湿度(Alemohammad等, 2018)、地表温度(Li等, 2019)等参量的通用降尺度方法. 同时, 神经网络(Wilby等, 1998; Cannon, 2011)、支持向量机(Ghosh, 2010)等机器学习模型也被广泛用于数值模拟数据的降尺度. 除了常规的降尺度方法, 图像处理领域的机器学习超分辨率技术也被引入用于提升地球系统模式输出的分辨率(Vandal等, 2017).

(3) 集成优化. 由于机理认知局限及参数化方案的差异，不同机理模型输出结果往往具有较大的不一致性，将不同模型输出进行综合是获得更可靠结果的有效途径. 在机器学习领域，集成学习通过结合多个机器学习器完成学习任务，可以达到模型间“博采众长”的效果，被广泛应用于遥感地表覆盖分类与制图的研究(杜培军和阿里木·赛买提, 2013). 同样，机器学习也可以实现对多个机理模型的集成优化，如图5所示. Monteleoni等(2011)基于隐马尔可夫模型对多个气候模型的预测结果进行集成，精度超过了原始最好的模型. 在此基础上，McQuade和Monteleoni(2012)进一步建立了更高空间分辨率的集成模型框架. Krasnopolsky和Lin(2012)利用神经网络进行多模型集成，使降水预报精度得到有效提升.

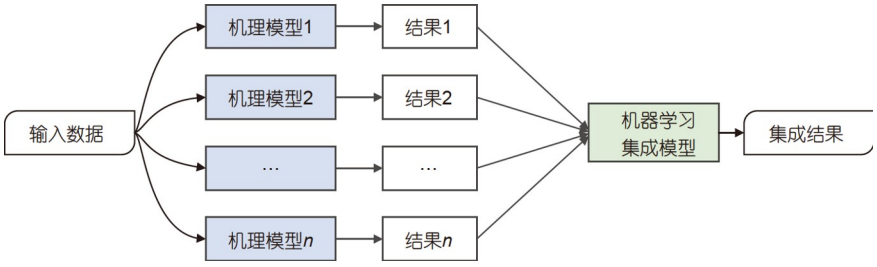

图 5　集成优化耦合方式

3.2　学习嵌入机理耦合范式

充分利用机理模型的物理可解释等建模优势，将学习模型嵌入到机理模型内部之中，并对原有的不确定子过程进行替代、调整或优化求解，是典型的以机理模型为主、学习模型为辅的耦合范式，也是当前机理-学习耦合研究的热点.

3.2.1　模型替代

模型替代即利用机器学习对机理模型的子过程进行替代的一种耦合方式，如图6所示. 在机理模型特别是动力学模型的建模过程中，一些子过程的空间尺度往往小于原有模型的网格尺度，以致难以用严格的物理模型进行直接建模，从而需要建立合适的参数化方案进行表达. 参数化就是对不能直接建模的物理过程进行间接表达的处理方案，是对复杂物理过程的近似或理想化表达(Stensrud, 2007). 因此，模型的参数化(Parameterization)与前述的参数优化(Parameter Optimization)有着本质区别. 在模型替代的耦合方式中, 应用最为广泛的就是利用机器学习模型替代机理模型中的参数化方案.

(1) 模型“仿真器”. 由于一些参数化方案的计算十分耗时，因此较为常用的一种替代方案就是以提升计算效率为目的, 通过对子模型输入-输出数据对的学习训练, 构建机理模型的机器学习“仿真器”, 使之具备与原有模型接近的精度以及更高的处理效率. Chevallier等(1999)将机器学习应用于新一代辐射传输模型的构建, 将多层感知器嵌入到整个物理建模过程之中, 用于替换从大气顶层到陆表的长波辐射，计算效率比传统带模式(band model)提升22倍，比逐线积分(line-by-line)模式提升10^6倍，该方法及其改进方案后续被业务化应用于欧洲中期天气预报中心的四维变分同化系统. 针对美国大气研究中心的CAM(Community Atmosphere Model)大气模型，Krasnopolsky等(2005)基于神经网络实现了对原有长波辐射参数化方法的仿真与替代, 并进一步应用于对流等过程的参数化处理(Krasnopolsky等, 2013), 计算效率可以比原有模型提升10~10^5倍(Krasnopolsky, 2020). Keller和Evans(2019)基于GEOS-Chem大气化学传输模式, 尝试用随机森林机器学习方法替换其中的化学积分器，形成了一套可行的替代方案, 为效率优化奠定了重要基础.

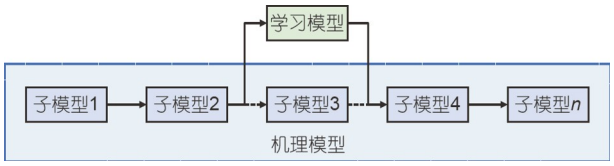

图 6　模型替代耦合方式

既然机器学习模型可以替代机理模型的部分子过程, 并达到与原有模型相近的计算精度, 人们自然好奇其是否可以替代更多子过程甚至整个复杂机理模型. Sargsyan等(2014)利用稀疏学习方法实现对陆面过程模型的仿真, 研究展现了一定的应用潜力. Krasnopolsky等(2009)基于美国国家环境预报中心的全球预报系统，尝试利用机器学习模型替代除辐射传输以外的所有子过程，发现并不是所有的输出都能够达到原有模型的水平. Dueben和Bauer (2018)利用深度学习构建了大气模型仿真器，针对区域的短期预测表现尚可，但长时序预测难以达到预期的效果. Scher和Messori (2019)研究表明在包含季节循环的大气模型中，机器学习替代整体机理模型仍然存在较大的挑战.

(2) 模型“增强器”. 如果存在足够的真实样本，机器学习替代方案还可以进一步提升估算精度. Bolton和Zanna(2019)在海洋参量模拟中，通过引入真实观测

数据与机器学习实现了模型的进一步优化，即使在仅有局部观测数据的条件下，也可以在全域尺度上提升模型的预测精度. Hunter等(2018)在河流参数模拟中，通过嵌入神经网络及简单的回归模型，有效提升了盐度的预测能力. Kraft等(2022)将神经网络模型嵌入到全球水文模型中，用于土壤湿度、地下水、雪等参数的模拟，获得了比机理模型更好的局部自适应性. 可见，如何充分利用高精度的地基观测、卫星遥感等数据，基于机器学习实现对不确定机理过程的替代，是实现模型提升的有效途径.

然而, 机器学习所需的训练样本经常难以获取. 为此, 可以利用更高分辨率的机理模型生成模拟数据, 将之作为“伪观测”数据进行学习模型的训练，然后将训练的模型应用于较低分辨率的机理模型中，如图7所示. 该方式已被广泛应用于大气模型的参数方案(Krasnopolsky等，2013；Schneider等，2017；Brenowitz和Bretherton，2018), 并被证明能够有效捕捉次格网尺度的时空信息, 获得比原有参数方案更高的精度, 甚至对极端事件都有较好的预测能力(Krasnopolsky等, 2009).

3.2.2 模型调整

如前所述，现有的全球和区域动力学模式通常都包含了复杂的参数化方案，从而导致了模型输出的不确定性(李新等，2007), 数据同化技术可以在模型的动力框架内，融合不同来源和不同分辨率的直接或间接观测，有效调整机理模型的运行轨迹，从而增强模型的预报精度及可预报性(李新等，2020，2021). 变分法、贝叶斯滤波是目前常用的两大类数据同化方法, 已有学者从数学上分析了数据同化和机器学习的理论等价性(Bonavita等，2021), 近年来如何将机器学习方法应用于数据同化已成为一个热点研究方向. 基于数据同化的模式，将机器学习方法嵌入到模型模拟的动力学框架之中，是实现机理-学习相互耦合的有效途径. 该模式与前述模型替代的区别在于并不直接替换模型原有的机理过程, 而是对其进行优化调整.

Hsieh和Tang(1998)较早提出在气象与海洋模式中利用机器学习进行数据同化的思想，研究学者利用神经网络(Härter和de Campos Velho, 2008)、支持向量机(Gilbert等, 2010)等机器学习模型进行了数据同化的理论探索, 并逐步应用于真实应用场景. 机器学习数据同化主要有三种方式: 第一, 利用机器学习对现有的同化

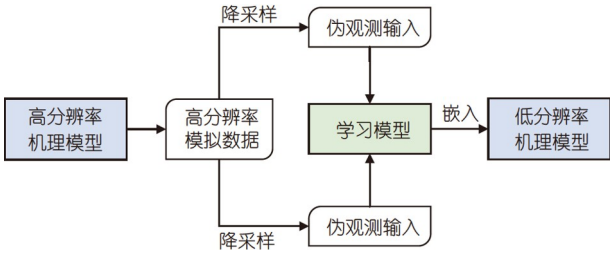

图 7　伪观测训练替代方案

算法进行模拟，其目的在于提升同化处理的效率，例如，在全球表面温度同化模拟研究中，达到相同精度的条件下神经网络方法比集合转换卡尔曼滤波效率提升274倍(Cintra等，2016)；第二，发展全新的机器学习同化方法，如Lu等(2018)利用神经网络同化算法有效提升了降水的预测精度；第三，将机器学习与现有数据同化方法结合，通过误差校正的方式提升模型的适用性(Bonavita和Laloyaux, 2020; Farchi等, 2021).

3.2.3 模型求解

在一些参量估算过程中，往往基于正向过程及相关物理机理建立最优化模型，并通过梯度下降迭代过程等进行模型求解，然而在具体的求解过程中，往往出现梯度不能计算或者即使可求解但计算量过大等问题, 此时可借助于机器学习进行模型的优化求解. 在理论研究方面, 机器学习被应用于求解偏微分方程, 该方向已在应用数学领域受到了广泛关注(Han等，2018). 在应用方面，Davis等(1993)在被动微波雪参数反演中，利用神经网络对散射模型进行训练学习，获得从参数到亮温的转换模型，并将之用于迭代求解算法；基于类似的求解思路，进一步反演了土壤湿度、近地气温、植被含水率等参数(Davis等，1995). Venkatakrishnan等(2013)发展了一种“即插即用”的机理-学习耦合方式，可以将机器学习模型嵌入到变分最优化迭代求解之中，用于SAR遥感数据重建(Alver等，2019)、多源数据融合(Dian等, 2021)等遥感应用中.

3.3 机理融进学习耦合范式

第三类耦合范式是将机理知识融进机器学习模型, 即以机器学习为主框架, 利用机理知识对学习过程进行约束引导，整个模型是一个“端到端”的计算方式. 根据机器学习模型中机理约束的施加位置及作用，可

分为输入变量约束、目标函数约束、模型结构约束等, 如图8所示(以神经网络为例).

3.3.1　输入变量约束

输入变量约束是指通过机理模型的计算或机理知识的引导, 为机器学习模型引入新的输入变量, 进而使学习过程更加符合特定的机理约束条件. 例如, 在Karpatne等(2017b)的研究中(如图9), 首先将驱动数据作为输入进行动力学机理模拟, 再将机理模拟的输出数据与原始驱动数据一同作为机器学习模型的输入变量, 此时机器学习的两组输入变量之间即存在相应的物理映射关系, 实验证明该耦合方式比纯数据驱动的模型具有更高的预测精度. 再如, Li等(2017)在遥感参量反演中, 在输入变量中引入了时空关联因子, 从而有效顾及了地理学第一定律, 对机器学习模型施加有效的时空地学约束.

3.3.2　目标函数约束

机器学习往往通过目标函数的最小化实现模型求解, 因此, 在目标函数中加入机理约束是一种直接易行并被广泛应用的融进方式(Kashinath等, 2021). 不失一般性, 可将机理约束神经网络的目标函数总结为如下基本形式(Karpatne等, 2017b; Willard等, 2020):

$$L = L_{\mathrm{d}}(x_{\mathrm{true}} - x_{\mathrm{pred}}) + \alpha R(w) + \beta L_{\mathrm{phy}}(x_{\mathrm{pred}}), \tag{1}$$

式中, 第一项L_{d}表征真实样本数据x_{true}与模型预测数据x_{pred}之间的监督误差, 可定义为误差平方和、绝对误差、交叉熵等形式; 第二项$R(w)$为通用正则化项, 具有压缩求解子集的作用, 其中w为模型的求解参数; 第三项$L_{\mathrm{phy}}(x_{\mathrm{pred}})$即为在通用正则化的基础上, 基于特定机理知识施加的约束, 以进一步缩小参数解的搜索空间, 更好地克服过拟合问题(Reichstein等, 2019); α、β为用于调节各项权重的超参数.

上式中$L_{\mathrm{phy}}(x_{\mathrm{pred}})$可以直接根据预测变量$x_{\mathrm{pred}}$的分布特征施加相应约束, 例如, Erichson等(2019)在目标函数中加入Lyapunov稳定性约束, 有效降低了海面温度预测的不确定性. 为了强化约束能力, 应用更为广泛的是引入与x_{pred}具有机理关联的变量z, 既可以是模型输入变量也可以是其他相关变量, 并以$L_{\mathrm{phy}}=\mathcal{L}(z, A, x_{\mathrm{pred}})$的形式进行约束, 其中$A$为机理关联模型, $\mathcal{L}$为某种惩罚函数. 例如, Karpatne等(2017b)在模拟湖水温度时引入密度变量, 充分利用温度与密度的物理关系方程, 并基于密度与深度的关系约束应用于L_{phy}的构建; 该方法后被进一步改进, 通过对输入-输出热通量的约束构建L_{phy}, 使得预测温度与湖水环境变化符合能量守恒定律(Read等, 2019; Jia等, 2021). Beucler等(2019)在模拟长波辐射过程中, 同时考虑了热量、质量、太阳辐射、地表辐射的守恒定律, 并在目标函数中施加相应的物理约束. 另外, 在遥感数据融合、降尺度等研究中, 可以将输入数据y与输出数据x_{pred}之间的正向模型用于L_{phy}的构建(Lin等, 2022), 如$L_{\mathrm{phy}} = \left\| y - Ax_{\mathrm{pred}} \right\|^2$, 即通过已知的关系矩阵$\mathbf{A}$对模型进行约束, 提升模型求解的保真度.

此外, 在一些具体的应用中, 也可以通过直接对L_{d}的改进实现对领域知识的引入. 例如, 当机器学习的目标变量无真实样本数据时, 也就无法直接构建目标函数, 然而如果存在与目标变量有确定机理关系的关联

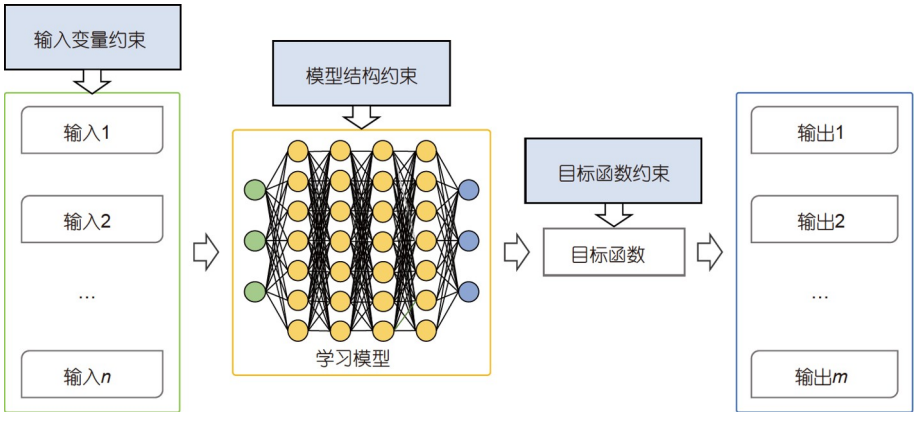

图 8　神经网络模型的机理约束

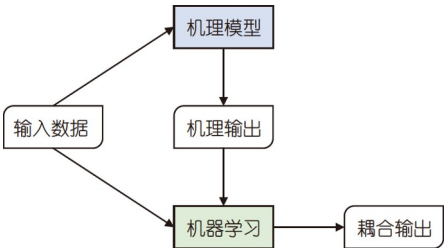

图 9　输入变量约束的一种形式

变量z, 则可以基于两种变量之间的机理关系间接构建目标函数, 如$L_{\mathrm{d}}=\left\| z-Bx_{\mathrm{pred}} \right\|^2$, 其中$B$为变量间的转换模型. De Bézenac等(2019)在海温估算时利用上述思路, 将运动场参数作为神经网络的目标变量, 并利用其与海温的物理关系建立能量函数, 实现二者的联合求解. 在不引入关联变量的条件下, Li等(2021)在定量遥感反演中建立时空地理加权约束函数$L_{\mathrm{d}}=\left\|w(x_{\mathrm{true}}-x_{\mathrm{pred}})\right\|^2$($w$为时空权值), 即通过顾及变量的自相关特征, 有效提升了模型的反演精度.

3.3.3　模型结构约束

机器学习求解过程往往是一个“黑箱”问题, 而另一种引入机理知识的方法就在“黑箱”中施加约束, 这就需要对机器学习内部结构和机理过程都有清晰的理解, 还需要找到它们之间的最佳结合点, 因此最具挑战性. Li等(2020)发展了时空地理加权学习方法, 对神经网络结构的模式层与求和层进行改进, 对加权求和节点、算术求和节点分别乘以相应的时空权值, 以充分考虑时空异质与时空相关的地学规律, 这与前述的目标函数时空约束方法(Li等, 2021)有异曲同工之妙, 但该方法难以应用于其他神经网络结构. Daw等(2020)直接在原有长短期记忆神经网络的后端添加一个激活函数, 并将激活函数输出用于表达湖水深度与密度的约束关系, 进而用于湖水温度的模拟. Beucler等(2019)同样将表达能量守恒的物理关系加到神经网络结构的后端(如图10), 形成了模型结构的约束方法, 并与基于目标函数约束方法进行了对比, 表明两种方法都可以有效改进长波辐射的模拟.

除了神经网络模型以外, 还有其他一些机器学习模型被用于模型结构的机理约束. 例如, 高斯过程回归是使用高斯过程先验对数据进行回归分析的一种非

图 10　神经网络模型的结构约束

参数模型, 对小样本数据十分有效, 并且能够分析预测的不确定性(Willard等, 2020). Camps-Valls等(2018)针对多输出的回归问题, 通过在高斯过程机器学习模型中引入微分方程, 从而对多变量之间的关系上施加物理约束, 并以叶面积指数与光合有效辐射为例, 验证了模型的有效性.

3.4　耦合范式的比较及混合应用

如上所述, 机理模型与学习模型的耦合包含级联、嵌入、融进三类基本范式, 其各有相应优势与限制条件. 级联范式的最大优势就是应用简单, 不需对两种模型的内部过程进行任何改动, 并且适用于多数的应用场景, 往往也能够获得明显的精度提升; 但是, 级联范式在理论上的突破有限, 缺乏对模型问题的根本性解决方案, 因此模型之间的互补优势难以得到充分发挥. 相对而言, 嵌入范式能够根据机理模型的不足进行针对性的改进, 由于其保持了机理模型的基本架构, 因此具备更强的物理可解释性, 适用于机理模型较为成熟的应用场景; 然而, 利用机器学习替代机理模型的子过程需要大量用于训练的中间变量, 而这些数据的获取往往比较困难, 成为一些应用中的限制条件. 融进范式保持了机器学习模型“端到端”的计算框架, 通过将机理知识融进学习模型实现高效的处理应用, 比较适合于机理模型不成熟同时又存在大量真实训练样本的场景; 然而对机器学习的“黑箱”模型结构进行修改往往十分困难, 对机理知识的引入程度会因此受到较大制约.

因此, 三种耦合范式并无绝对上的优劣之分, 而是针对不同的条件各有不同的适用场景, 并且在一些应用中它们还可以混合使用. 例如, Schneider等(2017)利用神经网络替代地球系统模式(ESM)中的参数化方案, 整体上属于学习嵌入机理的耦合范式, 然而, 其在神经网络的目标函数中融入了物理约束, 即进一步使用了机理融进学习的耦合范式. 再如, Read等(2019)在水温估算的研究中, 综合集成了三种耦合范式(Jia等, 2021):

首先应用级联范式, 基于机理模型生成模拟数据, 并进一步用于机器学习模型的预训练; 其次, 在精训练阶段将驱动数据与机理模型的输出一起作为机器学习模型的输入, 即融进范式; 再有, 在机器学习模型的能量函数中加入了物理约束项, 即嵌入范式. 通过不同耦合范式的联合应用, 可以更加充分地发挥机理模型与学习模型的互补优势.

4　主要问题与挑战方向

尽管机理-学习耦合在遥感反演和数值模拟领域都已经进行了前期探索, 并面向一些典型应用取得了可喜的进展, 但该方向整体上仍然处于较为初级的发展阶段, 在广度和深度上都亟待进一步发展. 在广度方面, 虽然针对大气数值模拟的研究相对较多, 但在陆面过程及水文模拟、遥感参量反演等方面的研究仍然较少, 亟需在各个方向上全面突破. 在深度方面, 如何在机理模型中嵌入更为稳健的学习过程, 如何在机器模型架构中融进更为复杂的机理知识, 仍然存在非常巨大的研究空间. 当前, 在地学大数据、人工智能迅猛发展的背景下, 机理模型与学习模型的耦合研究面临空前的机遇与挑战, 包括但不限于:

(1) 多源异类地学大数据的一体化学习与融合. 机器学习本身是一种数据驱动的计算模式, 因此机理-学习耦合很大程度上依赖于可用的参考数据. 当前, 虽然地基观测、遥感观测、数值模拟、社会感知等各类数据层出不穷, 但现有的模型耦合研究更多是针对单类或少类数据, 如何顾及多源异类数据在精度、尺度、时空连续性等方面的差异及其互补性, 开展一体化的机器学习建模与融合应用是重要的发展趋势(张良培和沈焕锋, 2016). 另外, 针对真实参考数据缺乏的问题, 如何充分利用多源多尺度观测与模拟数据, 通过迁移学习、主动学习等方式获得更为充足的训练样本, 是提升现有模型效能的有效途径.

(2) 机理过程的自适应学习替代机制. 利用机器学习替代机理模型中的不确定子过程, 因其保持了原有的物理过程机制、具有较强的物理可解释, 是具有极大潜力的一种耦合范式. 但在具体的执行过程中, 替代机理模型的哪个子过程? 什么情况下需要替代? 以及利用哪种机器模型进行替代? 这些都受多种因素的综合影响, 当前还主要依靠领域专家根据经验进行选定, 这就带来了一定的不确定性. 因此, 未来的发展趋势是构建机理过程的自适应学习替代机制, 即计算机根据机理模型运行情况和当前所具备的数据条件, 自动确定利用何种机器学习模型, 何时、何地替代何类机理子过程(von Rueden等, 2020).

(3) 面向地学领域的深度学习网络新架构. 深度学习作为当前最具代表性的机器学习方法, 正在成为地球表层参量估算及其机理-学习耦合的重要技术支撑. 然而, 现有方法主要还是基于通用神经网络架构, 如谷歌神经网络GoogLeNet(Szegedy等, 2015)、密集连接卷积网络DenseNet(Huang等, 2017)、残差网络ResNet(He等, 2016)、深度信念网络(Hinton等, 2006)等, 在原有基础上通过结构修改以引入机理或先验知识, 这就往往受固有网络结构限制而影响了模型的提升空间. 因此, 如何根据地学应用的独特性, 设计全新的深度神经网络架构, 是突破现有囿限的挑战性方向.

(4) 地学知识图谱与机器学习的耦合. 知识引导的机器学习方法是重要的发展方向, 而知识的表达具有多种形式, 知识图谱则是当前最为关注的方向之一(von Rueden等, 2023). 地学知识图谱就是以结构化的图模式, 将相关地学知识有效组织起来形成的一种知识体系, 表达了各类地理实体、概念及其之间的语义关系, 是可以被机器理解和计算的地学知识库和“推理机”(周成虎等, 2021). 众所周知, 符号主义、连接主义和行为主义是人工智能的三大流派, 而知识图谱、深度学习分别作为符号主义与连接主义的代表, 二者在地学领域的结合势必释放极大的应用潜能.

5　结语

针对地球表层特征参量的反演与模拟, 机理模型往往存在认知局限、欠定系统、计算负担等问题, 而学习模型经常在泛化性、迁移性、可解释性等方面存在不足. 机理模型与学习模型的耦合, 可以有效调整机理模型的“偏见”, 避免学习模型的“傲慢”(Chantry等, 2021), 是包括地学在内多个学科领域的重要关切方向. 本文面向特征参量的遥感反演与模型模拟, 搭建了机理级联学习、学习嵌入机理、机理融进学习及其混合应用的耦合范式框架, 结合具体应用实例系统总结了十种耦合模式, 进一步展望了一体化学习与融合、自适应学习替代、深度学习网络新架构、知识图

谱与深度学习耦合等挑战方向. 机理-学习耦合是“理性主义”与“经验主义”的结合, 将成为地球科学研究发展的“助推器”(Bergen等, 2019). 值得关注的是, 机理模型涉及严密的地学过程与物理推导, 学习模型需要建立复杂的信息传递机制, 迫切需要多学科交叉融合以突破机理-学习耦合的关键科学问题, 提升对地球表层参量的反演与模拟的精度与效率, 在地球系统科学研究、资源环境问题应对中展现更强的支撑能力.

致谢 本文的研究与写作受益于多位领域专家的启发、建议与帮助, 以及研究团队成员的交流与图文协助, 在此一并表示衷心感谢.

参考文献

陈春, 张志强, 林海. 2005. 地球模拟器及其模拟研究进展. 地球科学进展, 20: 1135–1142

陈发虎, 傅伯杰, 夏军, 吴铎, 吴绍洪, 张镱锂, 孙航, 刘禹, 方小敏, 秦伯强, 李新, 张廷军, 刘宝元, 董治宝, 侯书贵, 田立德, 徐柏青, 董广辉, 郑景云, 杨威, 王鑫, 李再军, 王飞, 胡振波, 王杰, 刘建宝, 陈建徽, 黄伟, 侯居峙, 蔡秋芳, 隆浩, 姜明, 胡亚鲜, 冯晓明, 莫兴国, 杨晓燕, 张东菊, 王秀红, 尹云鹤, 刘晓晨. 2019. 近70年来中国自然地理与生存环境基础研究的重要进展与展望. 中国科学: 地球科学, 49: 1659–1696

陈军, 刘万增, 武昊, Songnian L, 闫利. 2021. 智能化测绘的基本问题与发展方向. 测绘学报, 50: 995–1005

陈旻, 闾国年, 周成虎, 林珲, 马载阳, 乐松山, 温永宁, 张丰源, 王进, 朱之一, 许凯, 何元庆. 2021. 面向新时代地理学特征研究的地理建模与模拟系统发展与构建思考. 中国科学: 地球科学, 51: 1664–1680

程昌秀, 史培军, 宋长青, 高剑波. 2018. 地理大数据为地理复杂性研究提供新机遇. 地理学报, 73: 1397–1406

邓敏, 蔡建南, 杨文涛, 唐建波, 杨学习, 刘启亮, 石岩. 2020. 多模态地理大数据时空分析方法. 地球信息科学学报, 22: 41–56

杜培军, 阿里木·赛买提. 2013. 高分辨率遥感影像分类的多示例集成学习. 遥感学报, 17: 77–97

宫鹏. 2009. 遥感科学与技术中的一些前沿问题. 遥感学报, 13: 13–23

郭华东, 王力哲, 陈方, 梁栋. 2014. 科学大数据与数字地球. 科学通报, 59: 1047–1054

郭庆华, 金时超, 李敏, 杨秋丽, 徐可心, 巨袁臻, 张菁, 宣晶, 刘瑾, 苏艳军, 许强, 刘瑜. 2020. 深度学习在生态资源研究领域的应用: 理论、方法和挑战. 中国科学: 地球科学, 50: 1354–1373

郭仁忠, 林浩嘉, 贺彪, 赵志刚. 2020. 面向智慧城市的GIS框架. 武汉大学学报(信息科学版), 45: 1829–1835

胡斯勒图, 施建成, 李明, 王天星, 尚华哲, 雷永荟, 姬大彬, 闻建光, 阳坤, 陈良富. 2020. 基于卫星数据的地表下行短波辐射估算: 方法、进展及问题. 中国科学: 地球科学, 50: 887–902

黄昕, 李家艺, 杨杰, 张震, 李冬瑞, 刘小平. 2021. Landsat卫星观测下的30m全球不透水面年度动态与城市扩张模式(1972~2019). 中国科学: 地球科学, 51: 1894–1906

黎夏, 叶嘉安. 2005. 基于神经网络的元胞自动机及模拟复杂土地利用系统. 地理研究, 24: 19–27

李小文. 2005. 定量遥感的发展与创新. 河南大学学报(自然科学版), 35: 49–56

李新, 黄春林, 车涛, 晋锐, 王书功, 王介民, 高峰, 张述文, 邱崇践, 王澄海. 2007. 中国陆面数据同化系统研究的进展与前瞻. 自然科学进展, 17: 163–173

李新, 刘丰, 方苗. 2020. 模型与观测的和弦: 地球系统科学中的数据同化. 中国科学: 地球科学, 50: 1185–1194

李新, 马瀚青, 冉有华, 王旭峰, 朱高峰, 刘丰, 何洪林, 张臻, 黄春林. 2021. 陆地碳循环模型-数据融合: 前沿与挑战. 中国科学: 地球科学, 51: 1650–1663

李新, 郑东海, 冯敏, 陈发虎. 2022. 信息地理学: 信息革命重塑地理学. 中国科学: 地球科学, 52: 370–373

李召良, 段四波, 唐伯惠, 吴骅, 任华忠, 阎广建, 唐荣林, 冷佩. 2016. 热红外地表温度遥感反演方法研究进展. 遥感学报, 20: 899–920

梁顺林, 程洁, 贾坤, 江波, 刘强, 刘素红, 肖志强, 谢先红, 姚云军, 袁文平, 张晓通, 赵祥. 2016. 陆表定量遥感反演方法的发展新动态. 遥感学报, 20: 875–898

孟春雷, 戴永久. 2013. 城市陆面模式设计及检验. 大气科学, 37: 1297–1308

裴韬, 刘亚溪, 郭思慧, 舒华, 杜云艳, 马廷, 周成虎. 2019. 地理大数据挖掘的本质. 地理学报, 74: 586–598

邱晨辉. 2021. 我国首个地球系统模拟大科学装置启用. 中国青年报

冉有华, 李新, 程国栋, 南卓铜, 车金星, 盛煜, 吴青柏, 金会军, 罗栋梁, 唐志光, 吴小波. 2021. 2005~2015年青藏高原多年冻土稳定性制图. 中国科学: 地球科学, 51: 183–200

宋长青. 2016. 地理学研究范式的思考. 地理科学进展, 35: 1–3

吴志峰, 柴彦威, 党安荣, 龚建华, 高松, 乐阳, 李栋, 柳林, 刘行健, 刘瑜, 龙瀛, 陆锋, 秦承志, 王慧, 王鹏, 王伟, 甄峰. 2015. 地理学碰上“大数据”: 热反应与冷思考. 地理研究, 34: 2207–2221

张兵. 2018. 遥感大数据时代与智能信息提取. 武汉大学学报(信息科学版), 43: 1861–1871

张良培, 沈焕锋. 2016. 遥感数据融合的进展与前瞻. 遥感学报, 20: 1050–1061

张添, 黄春林, 沈焕锋. 2012. 土壤水分对土壤参数的敏感性及其参数优化方法研究. 地球科学进展, 27: 678–685

张正, 唐娉, 李宏益, 冯峥. 2016. 多源数据协同定量遥感产品生产系统的领域模型. 遥感学报, 20: 184–196

中国科学院地学部地球科学发展战略研究组. 2009. 21世纪中国地

球科学发展战略报告. 北京: 科学出版社
周成虎, 王华, 王成善, 侯增谦, 郑志明, 沈树忠, 成秋明, 冯志强, 王新兵, 闾海荣, 樊隽轩, 胡修棉, 侯明才, 诸云强. 2021. 大数据时代的地学知识图谱研究. 中国科学: 地球科学, 51: 1070–1079
Aires F, Prigent C, Rossow W B, Rothstein M. 2001. A new neural network approach including first guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature, and emissivities over land from satellite microwave observations. J Geophys Res, 106: 14887–14907
Alemohammad S H, Kolassa J, Prigent C, Aires F, Gentine P. 2018. Global downscaling of remotely sensed soil moisture using neural networks. Hydrol Earth Syst Sci, 22: 5341–5356
Alver M B, Saleem A, Cetin M. 2019. A novel plug-and-play SAR reconstruction framework using deep priors. Boston: Proceedings of the 2019 IEEE Radar Conference (RadarConf)
Anderson C. 2008. The end of theory: The data deluge makes the scientific method obsolete. Wired Magazine, 16: 16–07
Arnold J G, Srinivasan R, Muttiah R S, Williams J R. 1998. Large area hydrologic modeling and assessment part I: Model development. J Am Water Resour Assoc, 34: 73–89
Bauer P, Dueben P D, Hoefler T, Quintino T, Schulthess T C, Wedi N P. 2021. The digital revolution of Earth-system science. Nat Comput Sci, 1: 104–113
Beck H E, van Dijk A I J M, de Roo A, Miralles D G, McVicar T R, Schellekens J, Bruijnzeel L A. 2016. Global-scale regionalization of hydrologic model parameters. Water Resour Res, 52: 3599–3622
Bergen K J, Johnson P A, de Hoop M V, Beroza G C. 2019. Machine learning for data-driven discovery in solid Earth geoscience. Science, 363: aau0323
Beucler T, Rasp S, Pritchard M, Gentine P. 2019. Achieving conservation of energy in neural network emulators for climate modeling. arXiv preprint, arXiv:190606622. https://doi.org/10.48550/arXiv.1906.06622
Bolton T, Zanna L. 2019. Applications of deep learning to ocean data inference and subgrid parameterization. J Adv Model Earth Syst, 11: 376–399
Bonavita M, Geer A, Laloyaux P, Massart S, Chrust M. 2021. Data assimilation or machine learning? ECMWF Newsletter, No. 167
Bonavita M, Laloyaux P. 2020. Machine learning for model error inference and correction. J Adv Model Earth Syst, 12: e2020MS002232
Brenowitz N D, Bretherton C S. 2018. Prognostic validation of a neural network unified physics parameterization. Geophys Res Lett, 45: 6289–6298
Campos-Taberner M, García-Haro F J, Camps-Valls G, Grau-Muedra G, Nutini F, Crema A, Boschetti M. 2016. Multitemporal and multiresolution leaf area index retrieval for operational local rice crop monitoring. Remote Sens Environ, 187: 102–118
Camps-Valls G, Martino L, Svendsen D H, Campos-Taberner M, Muñoz-Marí J, Laparra V, Luengo D, García-Haro F J. 2018. Physics-aware Gaussian processes in remote sensing. Appl Soft Comput, 68: 69–82
Cannon A J. 2011. Quantile regression neural networks: Implementation in R and application to precipitation downscaling. Comput Geoscis, 37: 1277–1284
Chantry M, Christensen H, Dueben P, Palmer T. 2021. Opportunities and challenges for machine learning in weather and climate modelling: Hard, medium and soft AI. Phil Trans R Soc A, 379: 20200083
Chevallier F, Chéruy F, Scott N A, Chédin A. 1999. A neural network approach for a fast and accurate computation of a longwave radiative budget. J Appl Meteorol, 37: 1385–1397
Cintra R, De Campos Velho H, Cocke S. 2016. Tracking the model: Data assimilation by artificial neural network. Vancouver: Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN). 403–410
Davis D T, Zhengxiao Chen D T, Jenq-Neng Hwang D T, Tsang L, Njoku E. 1995. Solving inverse problems by Bayesian iterative inversion of a forward model with applications to parameter mapping using SMMR remote sensing data. IEEE Trans Geosci Remote Sens, 33: 1182–1193
Davis D T, Chen Z, Tsang L, Hwang J N, Chang A T C. 1993. Retrieval of snow parameters by iterative inversion of a neural network. IEEE Trans Geosci Remote Sens, 31: 842–852
Daw A, Thomas R Q, Carey C C, Read J S, Appling A P, Karpatne A. 2020. Physics-guided architecture (PGA) of neural networks for quantifying uncertainty in lake temperature modeling. In: Proceedings of the 2020 SIAM International Conference on Data Mining. 532–540
De Bézenac E, Pajot A, Gallinari P. 2019. Deep learning for physical processes: Incorporating prior scientific knowledge. J Stat Mech-Theory Exp, 2019: 124009
Dian R, Li S, Kang X. 2021. Regularizing hyperspectral and multispectral image fusion by CNN denoiser. IEEE Trans Neural Netw Learn Syst, 32: 1124–1135
Dueben P D, Bauer P. 2018. Challenges and design choices for global weather and climate models based on machine learning. Geosci Model Dev, 11: 3999–4009
Erichson N B, Muehlebach M, Mahoney M W. 2019. Physics-informed autoencoders for Lyapunov-stable fluid flow prediction. arXiv preprint, arXiv:190510866. https://doi.org/10.48550/arXiv.1905.10866

Farchi A, Laloyaux P, Bonavita M, Bocquet M. 2021. Using machine learning to correct model error in data assimilation and forecast applications. Q J R Meteorol Soc, 147: 3067–3084

Ganguly A R, Kodra E A, Agrawal A, Banerjee A, Boriah S, Chatterjee S, Chatterjee S, Choudhary A, Das D, Faghmous J, Ganguli P, Ghosh S, Hayhoe K, Hays C, Hendrix W, Fu Q, Kawale J, Kumar D, Kumar V, Liao W, Liess S, Mawalagedara R, Mithal V, Oglesby R, Salvi K, Snyder P K, Steinhaeuser K, Wang D, Wuebbles D. 2014. Toward enhanced understanding and projections of climate extremes using physics-guided data mining techniques. Nonlin Processes Geophys, 21: 777–795

Ghosh S. 2010. SVM-PGSL coupled approach for statistical downscaling to predict rainfall from GCM output. J Geophys Res, 115: D22102

Gilbert R C, Richman M B, Trafalis T B, Leslie L M. 2010. Machine learning methods for data assimilation. Comput Intell Architect Complex Eng Syst. New York: ASME Press. 105–112

Han J, Jentzen A, E W. 2018. Solving high-dimensional partial differential equations using deep learning. Proc Natl Acad Sci USA, 115: 8505–8510

Härter F P, de Campos Velho H F. 2008. New approach to applying neural network in nonlinear dynamic model. Appl Math Model, 32: 2621–2633

Härter F P, de Campos Velho H F. 2010. Multilayer perceptron neural network in a data assimilation scenario. Eng Appl Comput Fluid Mech, 4: 237–245

He K M, Zhang X Y, Ren S Q, Sun J. 2016. Deep residual learning for image recognition. Seattle, WA: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 770–778

Hinton G E, Osindero S, Teh Y W. 2006. A fast learning algorithm for deep belief nets. Neural Computation, 18: 1527–1554

Hsieh W W, Tang B. 1998. Applying neural network models to prediction and data analysis in meteorology and oceanography. Bull Amer Meteorol Soc, 79: 1855–1870

Huang G, Liu Z, Van Der Maaten L, Weinberger K Q. 2017. Densely connected convolutional networks. Honolulu: 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2261–2269

Hunter J M, Maier H R, Gibbs M S, Foale E R, Grosvenor N A, Harders N P, Kikuchi-Miller T C. 2018. Framework for developing hybrid process-driven, artificial neural network and regression models for salinity prediction in river systems. Hydrol Earth Syst Sci, 22: 2987–3006

Ivatt P D, Evans M J. 2020. Improving the prediction of an atmospheric chemistry transport model using gradient-boosted regression trees. Atmos Chem Phys, 20: 8063–8082

Jia X, Willard J, Karpatne A, Read J S, Zwart J A, Steinbach M, Kumar V. 2021. Physics-guided machine learning for scientific discovery: An application in simulating lake temperature profiles. ACM IMS Trans Data Sci, 2: 1–26

Karpatne A, Atluri G, Faghmous J H, Steinbach M, Banerjee A, Ganguly A, Shekhar S, Samatova N, Kumar V. 2017a. Theory-guided data science: A new paradigm for scientific discovery from data. IEEE Trans Knowl Data Eng, 29: 2318–2331

Karpatne A, Ebert-Uphoff I, Ravela S, Babaie H A, Kumar V. 2019. Machine learning for the geosciences: Challenges and opportunities. IEEE Trans Knowl Data Eng, 31: 1544–1554

Karpatne A, Watkins W, Read J, Kumar V. 2017b. Physics-guided neural networks (PGNN): An application in lake temperature modeling. arXiv preprint, arXiv:171011431. https://doi.org/10.48550/arXiv.1710.11431

Kashinath K, Mustafa M, Albert A, Wu J L, Jiang C, Esmaeilzadeh S, Azizzadenesheli K, Wang R, Chattopadhyay A, Singh A, Manepalli A, Chirila D, Yu R, Walters R, White B, Xiao H, Tchelepi H A, Marcus P, Anandkumar A, Hassanzadeh P, Prabhat P. 2021. Physics-informed machine learning: Case studies for weather and climate modelling. Phil Trans R Soc A, 379: 20200093

Keller C A, Evans M J. 2019. Application of random forest regression to the calculation of gas-phase chemistry within the GEOS-Chem chemistry model v10. Geosci Model Dev, 12: 1209–1225

Kraft B, Jung M, Körner M, Koirala S, Reichstein M. 2022. Towards hybrid modeling of the global hydrological cycle. Hydrol Earth Syst Sci, 26: 1579–1614

Krasnopolsky V. 2020. Using machine learning for model physics: An overview arXiv preprint, arXiv: 2002.00416. https://doi.org/10.48550/arXiv.2002.00416

Krasnopolsky V M, Fox-Rabinovitz M S, Belochitski A A. 2013. Using ensemble of neural networks to learn stochastic convection parameterizations for climate and numerical weather prediction models from data simulated by a cloud resolving model. Adv Artif Neural Syst, 2013: 1–13

Krasnopolsky V M, Fox-Rabinovitz M S, Chalikov D V. 2005. New approach to calculation of atmospheric model physics: Accurate and fast neural network emulation of longwave radiation in a climate model. Mon Weather Rev, 133: 1370–1383

Krasnopolsky V M, Lin Y. 2012. A neural network nonlinear multimodel ensemble to improve precipitation forecasts over continental US. Adv Meteorol, 2012: 1–11

Krasnopolsky V M, Lord S J, Moorthi S, Spindler T. 2009. How to deal with inhomogeneous outputs and high dimensionality of neural network emulations of model physics in numerical climate and weather prediction models. Atlanta: Proceedings of the International

Joint Conference on Neural Networks. 1668–1673
Lazer D, Kennedy R, King G, Vespignani A. 2014. The parable of google flu: Traps in big data analysis. Science, 343: 1203–1205
Li T, Shen H, Yuan Q, Zhang L. 2020. Geographically and temporally weighted neural networks for satellite-based mapping of ground-level $PM_{2.5}$. ISPRS J Photogrammetry Remote Sens, 167: 178–188
Li T, Shen H, Yuan Q, Zhang L. 2021. A locally weighted neural network constrained by global training for remote sensing estimation of $PM_{2.5}$. IEEE Trans Geosci Remote Sens, doi: 10.1109/TGRS.2021.3074569
Li T, Shen H, Yuan Q, Zhang X, Zhang L. 2017. Estimating ground-level $PM_{2.5}$ by fusing satellite and station observations: A geo-intelligent deep learning approach. Geophys Res Lett, 44: 11,985–11,993
Li W, Ni L, Li Z L, Duan S B, Wu H. 2019. Evaluation of machine learning algorithms in spatial downscaling of modis land surface temperature. IEEE J Sel Top Appl Earth Observ Remote Sens, 12: 2299–2307
Liang Z, Zou R, Chen X, Ren T, Su H, Liu Y. 2020. Simulate the forecast capacity of a complicated water quality model using the long short-term memory approach. J Hydrol, 581: 124432
Lin L, Li J, Shen H, Zhao L, Yuan Q, Li X. 2022. Low-resolution fully polarimetric SAR and high-resolution single-polarization SAR image fusion network. IEEE Trans Geosci Remote Sens, 60: 1–17
Lu J, Hu W, Zhang X. 2018. Precipitation data assimilation system based on a neural network and case-based reasoning system. Information, 9: 106
Mao K, Shi J, Li Z L, Tang H. 2007. An RM-NN algorithm for retrieving land surface temperature and emissivity from EOS/MODIS data. J Geophys Res, 112: D21102
McQuade S, Monteleoni C. 2012. Global climate model tracking using geospatial neighborhoods. AAAI, 26: 335–341
Monteleoni C, Schmidt G A, Saroha S, Asplund E. 2011. Tracking climate models. Statist Analy Data Min, 4: 372–392
Navares R, Aznarte J L. 2020. Predicting air quality with deep learning LSTM: Towards comprehensive models. Ecol Inf, 55: 101019
Noori N, Kalin L, Isik S. 2020. Water quality prediction using SWAT-ANN coupled approach. J Hydrol, 590: 125220
Petty T R, Dhingra P. 2018. Streamflow hydrology estimate using machine learning (SHEM). J Am Water Resour Assoc, 54: 55–68
Rasp S, Lerch S. 2018. Neural networks for postprocessing ensemble weather forecasts. Mon Weather Rev, 146: 3885–3900
Read J S, Jia X, Willard J, Appling A P, Zwart J A, Oliver S K, Karpatne A, Hansen G J A, Hanson P C, Watkins W, Steinbach M, Kumar V. 2019. Process-guided deep learning predictions of lake water temperature. Water Resour Res, 55: 9173–9190
Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat N. 2019. Deep learning and process understanding for data-driven Earth system science. Nature, 566: 195–204
von Rueden L, Mayer S, Beckh K, Georgiev B, Giesselbach S, Heese R, Kirsch B, Walczak M, Pfrommer J, Pick A, Ramamurthy R, Garcke J, Bauckhage C, Schuecker J. 2023. Informed machine learning—A taxonomy and survey of integrating prior knowledge into learning systems. IEEE Trans Knowledge Data Eng, 35: 614–633
von Rueden L, Mayer S, Sifa R, Bauckhage C, Garcke J. 2020. Combining machine learning and simulation to a hybrid modelling approach: Current and future directions. In: Berthold M, Feelders A, Krempl G, eds. Advances in Intelligent Data Analysis XVIII. Cham: Springer International Publishing. 548–560
Sargsyan K, Safta C, Najm H N, Debusschere B J, Ricciuto D, Thornton P. 2014. Dimensionality reduction for complex models via Bayesian compressive sensing. Int J Uncertain Quant, 4: 63–93
Sawada Y. 2020. Machine learning accelerates parameter optimization and uncertainty assessment of a land surface model. J Geophys Res-Atmos, 125: e2020JD032688
Scher S, Messori G. 2019. Weather and climate forecasting with neural networks: Using general circulation models (GCMs) with different complexity as a study ground. Geosci Model Dev, 12: 2797–2809
Schneider T, Lan S, Stuart A, Teixeira J. 2017. Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations. Geophys Res Lett, 44: 12,396–12,417
Shen H, Jiang M, Li J, Zhou C, Yuan Q, Zhang L. 2022. Coupling model- and data-driven methods for remote sensing image restoration and fusion: Improving physical interpretability. IEEE Geosci Remote Sens Mag, 10: 231–249
Shen H, Jiang Y, Li T, Cheng Q, Zeng C, Zhang L. 2020. Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data. Remote Sens Environ, 240: 111692
Shen H, Li T, Yuan Q, Zhang L. 2018. Estimating regional ground-level $PM_{2.5}$ directly from satellite top-of-atmosphere reflectance using deep belief networks. J Geophys Res-Atmos, 123: 13,875–13,886
Skamarock W, Klemp J, Dudhia J, Gill D, Barker D, Wang W, Powers J. 2005. A Description of the Advanced Research WRF Version 2. Technical Report. Report No. NCAR/TN 468+STR
Sønderby C K, Espeholt L, Heek J, Dehghani M, Oliver A, Salimans T, Agrawal S, Hickey J, Kalchbrenner N. 2020. Metnet: A neural weather model for precipitation forecasting. arXiv preprint, arXiv:200312140. https://doi.org/10.48550/arXiv.2003.12140

Stensrud D J. 2007. Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models. Cambridge: Cambridge University Press. 449

Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. 2015. Going Deeper with Convolutions. Boston: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 1–9

Trombetti M, Riaño D, Rubio M A, Cheng Y B, Ustin S L. 2008. Multi-temporal vegetation canopy water content retrieval and interpretation using artificial neural networks for the continental USA. Remote Sens Environ, 112: 203–215

Vandal T, Kodra E, Ganguly S, Michaelis A, Nemani R, Ganguly A R. 2017. DeepSD: Generating high resolution climate change projections through single image super-resolution. Halifax: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Part F129685: 1663–1672

Venkatakrishnan S V, Bouman C A, Wohlberg B. 2013. Plug-and-Play priors for model based reconstruction. Austin: 2013 IEEE Global Conference on Signal and Information Processing. 945–948

Wang F, Tian D, Lowe L, Kalin L, Lehrter J. 2021. Deep learning for daily precipitation and temperature downscaling. Water Res, 57: e2020WR029308

Wilby R L, Wigley T M L, Conway D, Jones P D, Hewitson B C, Main J, Wilks D S. 1998. Statistical downscaling of general circulation model output: A comparison of methods. Water Resour Res, 34: 2995–3008

Willard J, Jia X, Xu S, Steinbach M, Kumar V. 2020. Integrating physics-based modeling with machine learning: A survey. arXiv preprint, arXiv:200304919. https://doi.org/10.48550/arXiv.2003.04919

Witt C, Tong C, Zantedeschi V, Martini D, Kalaitzis F, Chantry M, Watson-Parris D, Bilinski P. 2021. RainBench: Towards global precipitation forecasting from satellite imagery. 35th AAAI Conference on Artificial Intelligence, AAAI 2021. 17A: 14902–14910

Wolanin A, Camps-Valls G, Gómez-Chova L, Mateo-García G, van der Tol C, Zhang Y, Guanter L. 2019. Estimating crop primary productivity with Sentinel-2 and Landsat 8 using machine learning methods trained with radiative transfer simulations. Remote Sens Environ, 225: 441–457

Xiao Q, Wang Y, Chang H H, Meng X, Geng G, Lyapustin A, Liu Y. 2017. Full-coverage high-resolution daily $PM_{2.5}$ estimation using MAIAC AOD in the Yangtze River Delta of China. Remote Sens Environ, 199: 437–446

Yuan Q, Shen H, Li T, Li Z, Li S, Jiang Y, Xu H, Tan W, Yang Q, Wang J, Gao J, Zhang L. 2020. Deep learning in environmental remote sensing: Achievements and challenges. Remote Sens Environ, 241: 111716

(责任编委: 李新)