

孙健,曹卓,李恒,等. 人工智能技术在数值天气预报中的应用. 应用气象学报, 2021, 32(1): 1-11.  
DOI: 10. 11898/1001-7313. 20210101

# 人工智能技术在数值天气预报中的应用

孙 健<sup>1)\*</sup> 曹 卓<sup>2)</sup> 李 恒<sup>2)</sup> 钱思萌<sup>3)</sup> 王 昕<sup>2)</sup> 闫力敏<sup>2)</sup> 薛 巍<sup>2)3)4)5)</sup>

<sup>1)</sup>(国家气象中心, 北京 100081)

<sup>2)</sup>(清华大学计算机科学与技术系, 北京 100084)

<sup>3)</sup>(国家超级计算无锡中心, 无锡 214072)

<sup>4)</sup>(清华大学地球系统科学系, 北京 100084)

<sup>5)</sup>(地球系统数值模拟教育部重点实验室, 北京 100084)

## 摘 要

当前,人工智能迎来第 3 次发展浪潮并在多个领域大数据分析中取得巨大成功,这为人工智能技术与数值天气预报结合提供了契机。已有大量研究尝试将人工智能技术用于数值天气预报的初值生成、预报和产品应用过程中,涉及观测资料预处理、资料同化、模式积分、后处理以及高性能计算,通过误差估计、参数估计和局部代理等手段使预报结果,得到改进且计算速度大幅提升,展示出良好的应用前景,一些神经网络模型也表现出纯数据驱动预报的可能性,在短时强对流天气、降水以及气候预测中已有较为理想的应用实例。然而,人工智能技术在数值天气预报中的应用与发展仍面临一些挑战,主要包括深度学习的弱解释性、不确定性分析以及两者的耦合等,除了应对这些挑战,未来两者的深度结合还需要在理论指导下的人工智能模型设计、高时空分辨率人工智能预报模型设计以及使用更多新型人工智能技术等方面深入探索。

**关键词:** 数值天气预报; 人工智能; 机器学习; 神经网络

## 引 言

人们的生产生活与天气变化息息相关,灾害性天气对能源<sup>[1]</sup>、农业<sup>[2-3]</sup>、运输<sup>[4-5]</sup>、军事<sup>[6]</sup>等方面具有重要影响,其中极端天气的影响尤为显著。据统计,2003—2017 年我国每年因气象灾害造成的直接经济损失约为 3000 亿元,且呈上升趋势<sup>[7]</sup>。因此,对未来天气状态有效、准确的预报带来的经济价值和社会意义巨大,远大于天气预报所需的基础设施投入。

为提高天气预报准确率,人们在探索天气预报技术的道路上从未停止脚步。纵观天气预报的发展历史,大致经历了民间技艺、单站预报、天气图预报和数值天气预报几个阶段<sup>[8]</sup>,这些阶段的演化也表现出天气预报技术内在的发展规律——使用更多的

数学、物理方法取代预报过程中的人工经验决策,不断丰富资料来源并开发更加有效的计算方法与计算工具,实现从纯人工经验的民间技艺到包含大量数学、物理形式化规律并建立在高性能计算技术之上的数值天气预报的转变。

目前数值天气预报发展面临着许多理论上和技术上的挑战,这些挑战来自初始误差、模式误差、天气系统演变的可预报性以及更高时空分辨率预报所需高计算复杂度等多个方面。同时,积累的海量数值天气预报信息在传统的数值预报产品中还未得到充分挖掘和拓展应用,等待人们提出行之有效的解决方法。20 世纪 80—90 年代,人们一致认为人工智能非常适合在大气科学中应用并具有广阔的前景,人工智能技术将会成为天气预报系统中必不可少的一部分,而不仅仅是一个辅助工具<sup>[9]</sup>。

近年人工智能迎来了第 3 次发展浪潮,以神经网络

2020-08-25 收到, 2020-11-02 收到再改稿。

资助项目: 国家重点研发计划(2017YFA0604500, 2016YFA0602100)

\* 邮箱: sunjian@cma.gov.cn

络为代表的深度学习方法在高维度、大数据量的众多复杂问题中表现优异,在理论和软硬件环境方面都取得了快速发展,我国也制定了《新一代人工智能发展规划》([http://www.gov.cn/zhengce/content/2017-07/20/content\\_5211996.htm](http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm))以抢占人工智能发展的重大战略机遇,这为人工智能技术在天气预报中的广泛应用提供了契机。因此,本文针对当前人工智能技术在数值天气预报中的应用与发展进行回顾,从数值天气预报全流程的角度,介绍目前预报过程中存在的问题以及正在探索的人工智能技术方案,包括观测资料预处理与资料同化、预报模式、后处理、高性能计算等多个方面,并对一些纯数据驱动的人工智能预报模型以及其中所采用的优化技术进行总结。

## 1 数值天气预报与人工智能技术

### 1.1 数值天气预报发展及主要技术

如今使用的数值天气预报技术已经历较长时间的发展。20 世纪初, Cleveland<sup>[10]</sup> 和 Bjerknes<sup>[11]</sup> 开始尝试将天气预报看作数学物理学中的初值问题,通过非线性偏微分方程组的形式描述旋转地球大气运动中的已知物理规律,在一定的初始条件下,求解方程组就可以得到未来一段时间的天气状态。由于无法求得方程组的解析解,因此只能将其离散化进行数值求解。偏微分方程数值求解是动力框架的核心,与描述辐射、云雨、边界层过程的物理参数化方案共同构成基本的大气环流模式。

为实现快速计算并使数值天气预报得以应用, Richardson<sup>[12]</sup> 首次对数值计算过程进行简化,降低方程组的复杂度,人们也开始尝试将计算任务交给计算机完成<sup>[13]</sup>,20 世纪 70 年代的超级计算机已有能力求解整套方程组<sup>[14]</sup>,高性能计算开始成为数值天气预报不可分割的一部分。

同时,人们注意到初值中的微小扰动会使天气预测结果产生巨大变化,因此如何减小初值误差成为提高预报准确性的关键问题之一,资料同化方法越来越多地用于获取更加精确的初值。另外,大气系统中的混沌特性以及数值预报中的初值误差、模式计算误差和物理参数化方案等导致的不确定性,使用单一模式预报结果不可避免地存在误差与不确定性,于是基于概率预报的集合预报方法逐渐被广泛使用<sup>[15]</sup>,资料同化与集合预报已成为提高预报质

量的重要方法。

如今观测资料预处理、资料同化、预报模式、集合预报和高性能计算等一起形成了高效、综合的数值天气预报业务系统<sup>[16-17]</sup>,涵盖初值生成、预报、产品应用等过程,基本预报流程如图 1 所示。按照预报时效,天气预报又分为临近预报(0~2 h)、短时预报(0~12 h)、短期天气预报(0~3 d)、中期天气预报(3~10 d)、延伸期预报(10~30 d)以及气候预测(30 d 以上)等多种业务模式<sup>[8,18]</sup>。数值天气预报业务系统已经孕育大量科学工程结晶,取得举世瞩目的成就,成为复杂工程系统的典范之一。经过 100 多年科学知识和先进技术的不断积累,数值天气预报能力取得稳定、持续的进步,充分显示出其内在的先进性<sup>[19]</sup>。与人工智能技术的融合也成为数值天气预报领域未来的热点研究内容之一,人工智能技术的加入将有助于海量观测资料和模拟结果的分析与深度应用,从而进一步提高数值天气预报产品服务质量。

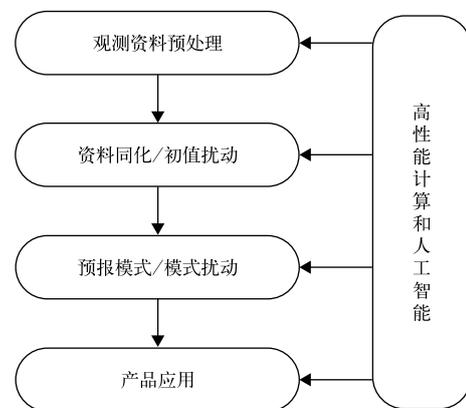


图 1 数值天气预报过程示意图

Fig. 1 Workflow of numerical weather prediction

### 1.2 人工智能技术

人工智能核心内容主要包括传统机器学习和深度学习两部分。

#### 1.2.1 传统机器学习

传统机器学习算法在少样本、低特征维度的情况下,能够获得与深度学习可比甚至更好的结果,且计算效率更高。传统机器学习中常见的问题包括分类、回归、聚类、降维、特征选择等<sup>[20]</sup>。常用分类或回归方法包括最近邻居法、决策树、集成方法、支持向量机等,聚类<sup>[21]</sup>方法包括 K 均值聚类(K-means)和层次聚类,常用降维方法包括主成分分析(PCA)和奇异值分解(SVD)等<sup>[22]</sup>,特征选择方法包括过滤

式、包裹式和嵌入式<sup>[23]</sup>。同时,机器学习方法也常用于数据的异常检测和缺失值重建。

### 1.2.2 深度学习和神经网络模型

深度学习主要用于解决传统机器学习方法在高维特征空间中拟合能力不足的问题,在处理图片、语音、文字等高维时空序列数据方面问题时也有出色表现。深度学习模型包括以深度前馈网络、卷积网络、递归网络、对抗生成网络和贝叶斯网络为代表的众多神经网络结构<sup>[20]</sup>。深度前馈网络是最基本的深度学习模型之一,其中各层神经元两两相连并通过反向传播算法实现参数更新,具有强大的非线性映射能力;卷积神经网络<sup>[24]</sup>引入卷积和池化操作,

通过系数交互、等变表示、参数共享等思想有效减少模型参数;递归神经网络<sup>[25]</sup>是专为处理序列数据提出的神经网络,通过将同一隐层神经元间进行连接的方式建立序列数据的前后关系;生成对抗网络<sup>[26]</sup>由生成器和判别器两部分组成,通过对抗学习,使生成器具有生成与训练数据分布相似样本的能力;贝叶斯网络基于贝叶斯推断采用嵌入隐变量生成样本空间的分布模型,在保持深度学习高拟合能力的同时提升泛化能力。

### 1.2.3 人工智能的发展方向

目前人工智能的发展方向主要体现在算法、算力和数据 3 个方面(图 2)。

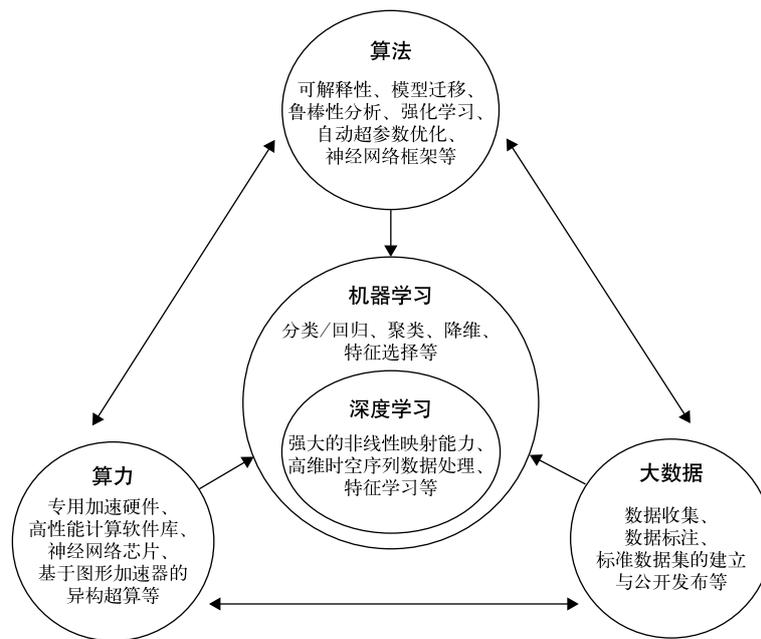


图 2 人工智能技术组成

Fig. 2 Components of artificial intelligence technology

在算法的开发方面,可解释性一直是深度学习算法追求的目标之一,模型具有可解释性便于知识提取、知识积累、知识重复利用,不可解释的模型也面临着稳定性、安全性等问题,同时,可解释性也能够使深度学习摆脱“黑盒”的标签,不再是毫无意义的暴力判定工具,目前可解释性研究主要包括模型结构或参数的可视化分析<sup>[27]</sup>、因果网络<sup>[28]</sup>开发等。另外,模型迁移和鲁棒性也是人工智能研究的热点,常用的迁移学习方法包括微调(Fine-tune)<sup>[29]</sup>、基于样本迁移、基于特征迁移和基于模型迁移等,现有的鲁棒性分析方法多采用样本攻击方法<sup>[30]</sup>。其他热点研究内容还包括自动超参数寻优(自动机器学习

框架(Auto-sklearn)<sup>[31]</sup>、神经网络结构搜索(NAS)<sup>[32]</sup>等)、强化学习<sup>[33]</sup>等。

人工智能软硬件也朝着更便捷、更高效的专用架构方向发展,软件环境主要包括 Python 中丰富的机器学习(Scikit-learn<sup>[34]</sup>等)和深度学习(Keras<sup>[35]</sup>、Tensorflow<sup>[36]</sup>、Pytorch<sup>[37]</sup>等)框架、CUDA 和 CuDNN 等图形处理器(GPU)高性能计算软件库。硬件环境包括神经网络芯片<sup>[38]</sup>、专用加速硬件(TPU 等)和通用加速器(GPU/FPGA)以及基于图形加速器的异构超级计算机(Summit<sup>[39]</sup>等)。软件框架的引入大大降低了人工智能技术应用门槛,高性能专用硬件能够有效提升训练和推理的计算速度,为超

大规模人工神经网络的采用和实时训练与推理等严苛场景应用奠定基础。

数据发展的主要内容是建立标准数据集,标准数据集的建立与公开发布有助于加快相关算法的研究与开发,如在图像识别问题中的手写数字数据集(mnist)、微小图像数据集(cifar10/cifar100)、大规模视觉识别图像数据集(imagenet)等。

### 1.3 人工智能与数值天气预报的联系及差异

人工智能和数值天气预报具有一定的相似之处,如人工智能与数值天气预报都是根据已知先验的信息预测未知,使用的数据都具有高维、海量、异构的特点,这些相似之处使人工智能技术能够很自然地融入数值天气预报中。

同时,人工智能与数值天气预报也存在较大差异:前者以数据为驱动,通过预测模型本身隐式表达数据中的规律;而后者以理论知识为驱动,注重将气象规律以数学物理公式的形式进行显式表达。两种方式各有优劣——数据驱动下规律的表达较为晦涩但能够自动、快速的实现,理论驱动下规律的表达更加清晰和准确但认识过程缓慢;又殊途同归——都

是以规律和知识获取为目标,只是方式和表达形式有所不同,如同统计学与数值计算的关系。这些差异的存在也为数值天气预报的发展提供新机遇,使两者各取所长、优势互补成为可能,如使用人工智能优化物理模型参数、替代物理模型子过程、代理模型等<sup>[40]</sup>,或利用物理理论指导、约束人工智能模型等<sup>[41-42]</sup>。

## 2 人工智能技术在数值天气预报中的应用

如今的数值天气预报过程主要涉及初值生成(观测资料预处理、资料同化)、预报(模式积分)、产品应用(后处理)和高性能计算(如图1所示),在当今人们对天气预报的需求中,每一部分都不可或缺,共同影响数值天气预报的精度与效率。然而,这些过程中仍然存在问题与挑战,有效解决和应对这些问题与挑战能够进一步提高数值预报性能,于是人们尝试使用人工智能技术帮助解决这些问题(表1)并取得一定进展。

表1 人工智能技术在数值天气预报中的应用

Table 1 Artificial intelligence applications to numerical weather prediction

功能	模块	人工智能技术	目标	效果
初值生成	观测资料处理及质量控制	贝叶斯方案、全卷积网络、极限学习机等	观测偏差校正 <sup>[43]</sup> 、雷达及卫星图像资料预处理 <sup>[44-45]</sup> 等	提高观测资料质量,优化高分辨率图像资料分割、资料填补等
	资料同化	随机森林、深度神经网络、支持向量机等	同化算法参数优化 <sup>[46]</sup> 、部分替代资料同化方法 <sup>[47]</sup> 、聚焦观测区域 <sup>[48]</sup> 等	提高同化质量,提高同化速度,更好利用高分辨率资料等
预报	模式积分	深度神经网络、卷积网络、随机森林等	模式代理 <sup>[49]</sup> 、替代物理过程参数化方案 <sup>[50-53]</sup> 、参数校正 <sup>[54-55]</sup> 等	提高模式计算速度,优化次网格物理过程的表示,提高参数校正效果与速度等
产品应用	后处理	随机森林、深度神经网络、卷积神经网络等	确定性及集合预报结果后处理 <sup>[56-58]</sup> 、替代集合预报 <sup>[59-60]</sup> 等	后处理偏差订正、质量更好、效率更高等

### 2.1 观测资料预处理与资料同化

由于观测资料预处理和资料同化共同影响初始条件的质量,且初始条件中的微小扰动会使模式结果产生较大差异,提高观测资料质量和改善资料同化方法显得尤为重要。大量观测资料凭借资料同化有效融入预测过程也使得人工智能在资料同化中的应用更具机遇。

在提高观测资料质量方面,Berry等<sup>[43]</sup>提出一种非参数贝叶斯方案对观测中的误差分布进行学习,并能够对其偏差进行纠正,与同化系统联合使用能够有效增强同化效果。另外,人工智能也可用于卫星图像等资料的预处理,如安捷等<sup>[44]</sup>使用全卷积

网络对气象卫星遥感图像资料中的云进行检测,实现高分辨率、大尺度、多通道遥感图像资料的云分割。Chang等<sup>[45]</sup>使用极限学习机填补中分辨率成像光谱仪反射资料中由于云量造成的空白。

在改善资料同化方法方面,Moosavi等<sup>[46]</sup>使用随机森林方法在时间和时空两种情境下对集合卡尔曼滤波资料同化方法中局部化函数的定位半径进行自适应调整,进一步提高同化质量。为提高同化速度,Cintra等<sup>[47]</sup>使用神经网络替代局部化集合转换卡尔曼滤波同化方法,在试验中得到相似的同化结果并表现出更高的计算效率。为更好地利用高空间分辨率资料进行同化,Lee等<sup>[48]</sup>使用支持向量机等

机器学习方法预先寻找预报中更感兴趣的资料区域,通过将高分辨率卫星放置在感兴趣区域以使资料同化产生更好结果。

人工智能有效改善了观测资料质量,并开始有机融入主流资料同化方法,但对于不确定性的量化与分析还需要更加深入的理论支持。

## 2.2 预报模式

数值天气预报模式涉及动力框架、物理过程参数化方案,其中模式数值积分过程最耗时,而人工智能技术推理具有高效性,使用人工智能模型替代模式(或部分模式)能够显著提高计算效率,这成为一个重要的研究方向。由于神经网络的“黑盒”特性,模式替代可以认为是代理模型(输入输出响应模型)的构建,如 Scher<sup>[49]</sup>使用深度神经网络学习简单的GCM模型输出,试验证明训练后的神经网络能够成功预测模型状态变化,从而取代动力学模型,使用神经网络预测的形式也大大提高计算速度,这一结果也展示了纯数据驱动天气预报的可能性。

物理参数化方案是当前模式中发展最为迅速的部分,也是不确定性最大的部分,使用神经网络模型替代与优化物理参数化方案的相关工作也被持续关注。目前物理参数化方案仍以单柱模型方式表示,随着分辨率的快速提升,次网格过程影响日益显著,三维物理方案的高计算量限制了其在模式中的应用,神经网络的高效拟合能力为同时降低计算开销提升计算精度创造了条件。Brenowitz等<sup>[50]</sup>针对传统物理过程参数化方案过于简化问题,探索在全球高空间分辨率资料下的神经网络参数化方法,使用4 km分辨率下的NG-Aqua(near-global Aqua-planet)仿真数据训练模型并对160 km网格视热源和视水汽汇进行预测,通过最小化多个时间步上的预测误差得到数值稳定的模型,在80 d的单柱试验中,相比CAM(Community Atmosphere Model)能够更好地匹配NG-Aqua结果。Pan等<sup>[51]</sup>将卷积神经网络作为统计降尺度方法替代现有降水参数化方案取得较理想的效果,并能够通过可视化分析对模型及结果进行合理解释。O'Gorman等<sup>[52]</sup>使用随机森林方法对湿对流自动参数化建模,仍能保证能量守恒和非负性降水量,带有机器学习湿对流参数化的大气环流模型能够稳定运行并准确捕获极端降水等特征。Rasp等<sup>[53]</sup>使用深度学习模型表示模式中的大气次网格过程,计算成本很小,突破了传统方法中的计算限制,且多年预报模拟结果

表现稳定,能够替代全球环流中传统的次网格参数化方法。Xu等<sup>[54]</sup>提出一种基于代理优化技术的碳循环模型参数校正方法,相比传统方法在计算准确度和计算效率两方面均得到一定提升。Wu等<sup>[55]</sup>针对CAM5提出一种自动参数估计方法并能对辐射平衡进行约束,结果显示其参数优化效果更好且能有效降低辐射不平衡。

同样,人工神经网络的物理参数化方案是否有能力表述包含复杂相变的云、雨、雪、气溶胶等过程与边界层的湍流过程,以及与动力框架的可靠耦合仍值得深入探究。

## 2.3 集合预报后处理与集合代理

集合预报是解决单一数值预报结果不确定问题的有效途径,从统计学角度给出概率预报结果并对预报的可信度进行量化。集合预报中的问题主要体现在集合预报结果后处理和集合预报效率等方面。在集合预报结果后处理方面,Burke<sup>[56]</sup>使用随机森林方法对高分辨率集合预报结果进行订正,改善冰雹预报效果,订正后的预报结果具有更高的可靠性和更小的模型偏差,同时由于不需要复杂物理假设计算效率也更高。Taillardat<sup>[57]</sup>使用分位数回归森林对集合预报结果进行后处理,其中分位数通过集合成员之外的其他预测因子估计得到,消除人为假设订正变量的局限性,提高订正效果。Rasp等<sup>[58]</sup>提出一种基于神经网络的集合预报后处理方案,通过数据驱动方式自动学习任意预测变量和预测分布参数间的非线性关系,避免预先指定关系函数的局限性。试验结果证明该方法在计算上更经济,且更易推广到其他统计后处理和预测问题中。

为有效提高集合预报效率,Scher等<sup>[59]</sup>尝试使用基于卷积神经网络的深度学习替代集合预报方法,并能够对给定初始状态下模式预报结果的不确定性进行估计,相比传统集合方法计算代价更小,其预报性能的限制在于可用的训练样本较少。Sonderby等<sup>[60]</sup>使用深度神经网络进行降水预报时直接输出概率预报结果,有效替代集合概率预报方法且预报速度更快。

人工智能技术被认为是一种拟合高维非线性样本空间概率分布,量化不确定性的有效手段,帮助或替代传统的集合预报是非常有潜力的研究方向,并开展初步工作。通过对于初始场的学习,开展集合方案的设计值得进一步关注。将资料同化与集合预报结合起来进行人工智能创新更值得期待。

## 2.4 人工智能预报模型

一些研究者更加大胆地尝试利用人工智能模型直接进行天气预报,独立于传统的数值预报过程,其预报流程如图 3 所示。这些人工智能模型以观测资料为输入并直接输出预报结果,能够自动捕捉输入资料中的时间和空间特征并作出快速推理,其结果与数值天气预报模式相比具有相近甚至更高的预报精度,实际预测时(推理过程)计算速度也具有明显优势。

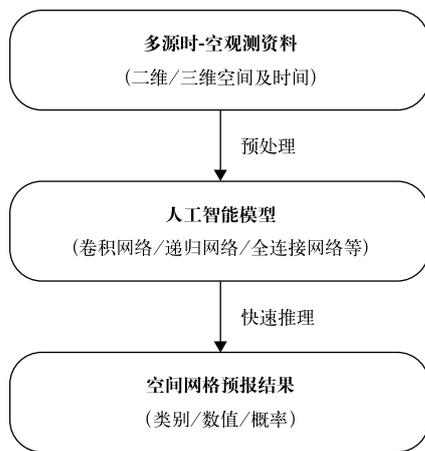


图 3 基于人工智能模型的天气预报流程

Fig. 3 Weather prediction workflow based on artificial intelligence models

Ham 等<sup>[61]</sup>采用深度学习方法能够对 ENSO 事件进行长达 1 年半的预测,且预测效果远高于目前最先进的热带气候变化动态预报系统(SINTEX-F 等),证明卷积神经网络是预测 ENSO 事件的有力工具。在降水预报方面,Google 公司的 Sonderby 等<sup>[60]</sup>提出了深度学习降水预报模型 MetNet,实现 1 km 空间分辨率和 2 min 时间分辨率下全美范围内提前 7~8 h 的降水预报,超过 NOAA 使用的大气模型结果,且计算时间从几十分钟或几小时缩减至几秒。在极端天气预报方面,Zhou 等<sup>[62]</sup>提出一种短时暴雨、冰雹、对流阵风和雷暴等强对流天气的深度学习预报方法,使用温度、气压、湿度、风以及数十个对流物理参数作为预报因子,建立 6 层卷积神经网络作为预测模型,其结果优于其他传统方法,该深度学习预报模式目前已在国家气象局国家气象中心应用。

虽然人工智能模型应用取得了初步成果,但这些模型多为特征变化显著的短时天气预报或特征变化缓慢的长期气候信号的预测,对于物理现象复杂

且特征多变的天气过程预报,人工智能模型还未出现较好的应用案例,人工智能模型在解决高维非线性天气问题中的表现还需要进一步验证。

## 3 人工智能技术的应用

近年来,人工智能应用快速发展,针对不同领域和不同需求的应用发展了大量技术手段。

### 3.1 大规模神经网络计算

从多源观测大数据中学习知识,实现高维、复杂和多变量的天气过程特征识别和状态预测,需要借助大规模神经网络实现。与传统数值天气预报需要依赖超级计算机支撑一样,大规模神经网络计算也需要强大的算力支撑,而预报时效进一步提升了算力需求。

Kurth 等<sup>[63]</sup>在超算 Piz-Daint 和 Summit 上使用 Tiramisu 和 DeepLab v3+ 深度学习网络模型对高时空分辨率下的极端天气进行模式识别与分割,在识别极端天气位置的同时对边界进行划分。另外,通过对软件框架和训练算法的改进,DeepLab v3+ 网络可使用到 27360 个 V100 GPU,持续吞吐为 325.8 PFLOPS 且并行计算效率为 90.7%,通过利用 FP16 张量核,半精度下的 DeepLab v3+ 网络峰值和持续吞吐分别高达 1.13 EFLOPS 和 999 PFLOPS,这一应用获得 2018 年戈登贝尔奖,展示人工智能技术在气象领域中进行大规模计算的可能性。Rojek<sup>[64]</sup>针对改进的随机森林方法提出一种混合精度算法,在基于 GPU 的超算中心进行天气预报模拟时,该算法能够提供与双精度相当的计算精度,且计算能耗降低 36%。

### 3.2 特征分析与可解释性

特征分析与可解释性有助于人们理解人工智能模型的推理过程,便于对预报模型进行优化。Manandhar 等<sup>[65]</sup>通过特征选择方法找到对降水具有较大影响的特征,基于这些特征使用机器学习方法对降水量进行预测,获得较高预测准确率的同时有效降低虚警率。Gagne 等<sup>[66]</sup>使用卷积网络基于上层空气动力场和热力场预测严重冰雹发生概率,通过对卷积网络特征重要性的解释与分析,发现该网络综合了对冰雹生成过程的理解,预测结果也优于其他模型。

### 3.3 样本权重差异化与注意力机制

人工智能模型在推理天气情况时使用大量时间

或空间相关数据,这些数据可以被赋予不同权重用以表示其对预报点影响的大小。Karevan 等<sup>[67]</sup>利用长短期记忆网络建立数据驱动的天气预报模型,考虑样本与预报点距离的影响,将预报点附近的样本赋予更大权重,所提出的传导长短期记忆网络在预测任务中表现出较好的性能。Qiu 等<sup>[68]</sup>提出一种多任务卷积神经网络用于降水量短期预测,考虑到观测的不完整性以及临近点影响,该模型通过多任务处理观测点的时间序列数据并利用多个观测点间的相关性进行预测,结果优于包括欧洲中期天气预报中心模型在内的一系列基线模型。Yuan 等<sup>[69]</sup>提出一种基于双因素注意力机制的多变量编码-解码预报模型,从时间信息和先验知识推断两方面学习注意力权重,在公共天气预报数据集上采用双因素注意力机制方法时预报精度得到很大提高。

### 3.4 基于物理机制的定制化目标函数

目标函数确定对人工智能神经网络质量的影响至关重要。Prasetya 等<sup>[70]</sup>从计算机视觉角度考虑降水预报中的雷达回波外推任务,将图像质量评估方法中的结构相似性指标和均方误差、平均绝对误差结合并作为神经网络训练目标,使雷达图像预测质量得到有效提高。Tan 等<sup>[71]</sup>提出一种基于分层卷积长短期记忆网络的深度学习模型,并使用一种新的预报员损失函数预报未来的卫星云图,所提出的神经网络具有更好的预报性能,预报员损失函数也能更好地保留实际大气状况的不确定性。

## 4 面临的挑战

人工智能技术在数值天气预报领域展现出巨大潜力,主要体现在两个方面:人工智能技术尝试在数值天气预报的核心部分发挥作用,如对观测资料预处理、资料同化、预报模式、集合预报、后处理等进行优化和改善;纯数据驱动的端到端人工智能预报模型开发。这些模型也表现出替代传统数值模式的可能性,同时大量人工智能技术应用于天气预报并对其发展起积极作用。然而,目前人工智能技术在数值天气预报中也面临一些挑战:

①弱解释性问题。传统数值天气预报发展一直十分注重理论知识的积累,与人工智能在驱动方式上有很大差别,人工智能多以统计为基础,对规律的解释较晦涩,解释性不足,基于人工智能的知识发现也较困难,在一定程度上限制了其在行业内的广泛

认可及广泛应用。

②不确定性分析问题。如今的数值天气预报十分注重不确定性分析,如何使用人工智能技术有效表述、设计初值及模式不确定性,需特别考虑。

③人工智能与现有数值天气预报技术耦合问题。现有的人工智能应用表现出局部优化、局部替代、全局替代等多种耦合模式,然而对于耦合系统的长期稳定运行还缺乏深刻的理论分析和保障。人工智能长于对于样本分布的估计,传统数值预报技术优于确定性规律下的精确计算,这对耦合设计提出了挑战。另外,数值天气预报和人工智能在代码、软硬件环境方面的需求也存在一定差异,需妥善解决。

④理论知识指导下的人工智能模型设计。数值天气预报领域至今已积累大量的数学和物理理论,这些理论应被人工智能模型有效利用,如何合理使用这些理论指导人工智能模型设计值得进一步讨论。

⑤人工智能预报模型在高时空分辨率天气预报问题中的可行性。尽管现有的人工智能预报模型应用取得了初步成果,但在高时空分辨率条件下强非线性问题中使用纯数据驱动的人工智能预报模型的可行性仍需验证。同时,人工智能预报模型在训练过程中计算量和数据访问量也大幅提升,这本身也对大规模高性能计算机软硬件平台提出了新的要求。

## 5 展望

人工智能技术在数值天气预报应用中取得成功的同时仍面临挑战,人工智能技术可重点从以下几个方面发展:

1) 在数值天气预报业务中,初值生成过程特别是资料同化方法对预报质量影响尤为重要,而受限于计算压力难以有效利用海量高时空分辨率气象资料,这或许是目前人工智能最为有效的着力点之一。在预报模式、后处理等过程中合理使用代理模型,以及构建端到端的数据驱动模型,能够不失准确性又显著提高预报时效,为大规模计算任务提供高效方案。

2) 在人工智能模型知识积累和再开发方面,贝叶斯网络、图神经网络和因果网络的发展在一定程度上有助于建立更为全面深刻的特征工程,支持理论知识发现与知识积累;在线学习与迁移学习技术能够支持人工智能预报模型的积累与增量开发过

程,提高历史资料与模型利用率;自动网络结构与超参数寻优方法能够实现机器学习和深度学习模型中超参数的自动化快速搜索,加快人工智能应用开发。

3) 针对人工智能模型的不确定性,采用贝叶斯推断技术生成当前气象状态的分布特征可能成为更全面、更高效的不确定性量化替代方案;基于贝叶斯网络嵌入隐藏变量特征的学习,基于注意力机制对气象要素间长程依赖关系的识别,可能为系统不确定性分析提供更有洞察力的特征抽取和机理发现。这方面潜力的挖掘需要从数学层面找到与传统分析方法的联系,从而建立综合的不确定性分析方法,同时,应使用合理的分析与评价方法改善人工智能模型的鲁棒性以减小不确定性,如将数学或物理理论作为先验知识用于神经网络结构设计、定制损失函数、控制模型输出等,从而使结果满足一定物理约束条件并具有更小的不确定性。

4) 数值天气预报模式与人工智能模型的耦合需要在耦合方案以及算法的数值稳定性和精度方面进行探究。标准工作流和框架的开发将有助于人工智能耦合模型的开发,如实现 Fortran, MPI, NetCDF, NCL 一系列基础语言和工具与 PyTorch, TensorFlow 等神经网络框架有效连接,以便代码耦合与继承。

5) 为实现高时空分辨率人工智能预报模型开发,超大规模神经网络的针对性研究或能满足精细化预报的需求,超大规模网络实现高斯随机过程的收敛性理论和实践突破,可有效支撑高度非线性模式的同化和集合设计。另外,人工智能分布式学习框架和人工智能超算的发展也为大规模人工智能技术应用提供完整和高效的解决方案。

数值天气预报融大数据与大计算于一体,成为复杂工程系统的典范之一。如今,人工智能迎来了第3次发展浪潮,为其在数值天气预报中的深入应用提供了契机。从现有研究看,人工智能已开始数值天气预报中发挥作用,体现在观测资料预处理、资料同化、模式预报、后处理等多个方面,包括模式代理与参数化方案优化、同化算法优化与替代、观测资料预处理与预报结果后处理等。纯数据驱动的人工智能预报模型也已开发并投入使用,表现出人工智能在天气预报问题中的巨大潜力。人工智能与数值天气预报的广泛结合还有很长的路要走,需要气象领域专家与人工智能、高性能计算等计算机领域专家携手开展深度合作研究。

## 参考文献

- [1] 段海来,千怀遂.广州市城市电力消费对气候变化的响应.应用气象学报,2009,20(1):80-87.  
Duan Hailai, Qian Huaisui. Responses of the electric power consumption to climate change in Guangzhou City. *J Appl Meteor Sci*, 2009, 20(1): 80-87.
- [2] 郭建平.农业气象灾害监测预测技术研究进展.应用气象学报,2016,27(5):620-630.  
Guo Jianping. Research progress on agricultural meteorological disaster monitoring and forecasting. *J Appl Meteor Sci*, 2016, 27(5): 620-630.
- [3] 王纯枝,霍治国,张蕾,等.北方地区小麦蚜虫气象适宜度预报模型构建.应用气象学报,2020,31(3):280-289.  
Wang Chunzhi, Huo Zhiguo, Zhang Lei, et al. Construction of forecasting model of meteorological suitability for wheat aphids in the Northern China. *J Appl Meteor Sci*, 2020, 31(3): 280-289.
- [4] 周雨,刘志萍,张国平.鹰厦铁路降水诱发地质灾害概率预报模型及应用.应用气象学报,2015,26(6):743-749.  
Zhou Yu, Liu Zhiping, Zhang Guoping. Probability forecasting model of geological disaster along the Yingxia Railway induced by precipitation with its application. *J Appl Meteor Sci*, 2015, 26(6): 743-749.
- [5] 侯英雨,张蕾,吴门新,等.国家级现代农业气象业务技术进展.应用气象学报,2018,29(6):641-656.  
Hou Yingyu, Zhang Lei, Wu Menxin, et al. Advances of modern agrometeorological service and technology in China. *J Appl Meteor Sci*, 2018, 29(6): 641-656.
- [6] 高太长,刘磊,赵世军,等.全天空测云技术现状及进展.应用气象学报,2010,21(1):101-109.  
Gao Taichang, Liu Lei, Zhao Shijun, et al. The actuality and progress of whole sky cloud sounding techniques. *J Appl Meteor Sci*, 2010, 21(1): 101-109.
- [7] 翟潘毛,刘静.极端天气/气候事件和灾害预防与减灾的全球变暖背景. *Engineering Sciences*, 2012, 14(9): 55-63.
- [8] 穆穆,陈博宇,周非凡,等.气象预报的方法与不确定性.气象,2011,37(1):1-13.  
Mu Mu, Chen Boyu, Zhou Feifan, et al. Methods and uncertainties of meteorological forecast. *Meteorological Monthly*, 2011, 37(1): 1-13.
- [9] 曾晓梅.国外人工智能技术在天气预报中的应用综述.气象科技,1999,27(1):4-10.  
Zeng Xiaomei. Application of artificial intelligence technology in weather forecast abroad. *Meteorological Science and Technology*, 1999, 27(1): 4-10.
- [10] Cleveland A. The physical basis of long-range weather forecasts. *Mon Wea Rev*, 1901, 29(12): 551.
- [11] Bjerknes V. Das Problem der Wettervorhersage betrachtet vom Standpunkt der Mechanik und Physik. *Meteorol Z*, 1904, 21: 1-7.
- [12] Richardson L F. Weather Prediction by Numerical Process. Cam-

- bridge; Cambridge University Press, 1922.
- [13] Charney J G, Fjoertoft R, Von Neumann J. Numerical integration of the barotropic vorticity equation. *Tellus*, 1950, 2: 237-254.
- [14] Lynch P. The origins of computer weather prediction and climate modeling. *J Comput Phys*, 2008, 227: 3431-3444.
- [15] 李泽椿, 陈德辉. 国家气象中心集合数值预报业务系统的发展及应用. *应用气象学报*, 2002, 13(1): 1-15.  
Li Zechun, Chen Dehui. The development and application of the operational ensemble prediction system at National Meteorological Center. *J Appl Meteor Sci*, 2002, 13(1): 1-15.
- [16] 沈学顺, 苏勇, 胡江林, 等. GRAPES\_GFS 全球中期预报系统的研发和业务化. *应用气象学报*, 2017, 28(1): 1-10.  
Shen Xueshun, Su Yong, Hu Jianglin, et al. Development and operation transformation of GRAPES global middle-range forecast system. *J Appl Meteor Sci*, 2017, 28(1): 1-10.
- [17] 贺雅楠, 高嵩, 薛峰, 等. 基于 MICAPS4 的智能网格预报平台设计与实现. *应用气象学报*, 2018, 29(1): 13-24.  
He Yanan, Gao Song, Xue Feng, et al. Design and implementation of intelligent grid forecasting platform based on MICAPS4. *J Appl Meteor Sci*, 2018, 29(1): 13-24.
- [18] 李泽椿, 毕宝贵, 金荣花, 等. 近 10 年中国现代天气预报的发展与应用. *气象学报*, 2014, 72(6): 1069-1078.  
Li Zechun, Bi Baogui, Jin Ronghua, et al. The development and application of the modern weather forecast in China for the recent 10 years. *Acta Meteorologica Sinica*, 2014, 72(6): 1069-1078.
- [19] Bauer P, Thorpe A, Brunet G. The quiet revolution of numerical weather prediction. *Nature*, 2015, 525(7567): 47-55.
- [20] Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge: MIT Press, 2016.
- [21] Jain A K. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 2009, 31(8): 651-666.
- [22] Yang J, Zhang D, Frangi A F, et al. Two-dimensional PCA: A new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(1): 131-137.
- [23] 周志华. 机器学习. 北京: 清华大学出版社, 2016.  
Zhou Zhihua. Machine Learning. Beijing: Tsinghua University Press, 2016.
- [24] Hadji I, Wildes R P. What Do We Understand About Convolutional Networks? Preprint at <https://arxiv.org/abs/1803.08834>, 2018: 1-94.
- [25] Graves A. Supervised sequence labelling with recurrent neural networks. *Studies in Computational Intelligence*, 2012, 385: 1-131.
- [26] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets//Proceedings of the 27th International Conference on Neural Information Processing Systems, 2014: 2672-2680.
- [27] Qin Z, Yu F, Liu C, et al. How convolutional neural networks see the world—A survey of convolutional neural network visualization methods. *Mathematical Foundations of Computing*, 2018, 1(2): 149-180.
- [28] Pearl J, Mackenzie D. The Book of Why. London: Allen Lane, 2019.
- [29] Yosinski J, Clune J, Bengio Y, et al. How Transferable are Features in Deep Neural Networks?//Proceedings of the 27th International Conference on Neural Information Processing Systems, 2014: 3320-3328.
- [30] Akhtar N M A. Threat of adversarial attacks on deep learning in computer vision: A Survey. *IEEE Access*, 2018, 6: 14410-14430.
- [31] Feurer M, Klein A, Eggensperger K, et al. Efficient and robust automated machine learning. *Advances in Neural Information Processing Systems*, 2016, 28: 2944-2952.
- [32] Jin H, Song Q, Hu X. Auto-Keras: An Efficient Neural Architecture Search System//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019: 1946-1956.
- [33] Francois-Lavet V, Henderson P, Islam R, et al. An introduction to deep reinforcement learning. *Foundations and Trends in Machine Learning*, 2018, 11(3-4), DOI: 10.1561/22000000071.
- [34] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 2011, 12: 2825-2830.
- [35] Chollet F. Keras(2020-04-28)[2020-06-20]. <https://keras.io>, 2020.
- [36] Abadi M, Barham P, Chen Jianmin, et al. TensorFlow: A System for Large-scale Machine Learning // Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation, 2016: 265-283.
- [37] Paszke A, Gross S, Chintala S, et al. Automatic Differentiation in PyTorch//31st Conference on Neural Information Processing Systems, 2017: 1-4.
- [38] Yin S, Ouyang P, Tang S, et al. A high energy efficient reconfigurable hybrid neural network processor for deep learning applications. *IEEE Journal of Solid-State Circuits*, 2018, 53(4): 968-982.
- [39] Vazhkudai S, Supinski B R, Bland A S. The Design, Deployment, and Evaluation of the CORAL Pre-exascale Systems//The International Conference for High Performance Computing, Networking, Storage, and Analysis, 2018: 661-672.
- [40] Reichstein M, Camps-Valls G, Stevens B, et al. Deep learning and process understanding for data-driven earth system science. *Nature*, 2019, 566(7743): 195-204.
- [41] Karpatne A, Watkins W, Read J, et al. Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling. Preprint at <https://arxiv.org/abs/1710.11431v2>, 2017: 1-11.
- [42] Karpatne A, Atluri G, Faghmous J, et al. Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge & Data Engineering*, 2017, 29(10): 2318-2331.
- [43] Berry T, Harlim J. Correcting biased observation model error in data assimilation. *Mon Wea Re*, 2017, 145(7): 2833-2853.
- [44] 安捷, 马尽文. 基于全卷积网络的遥感图像自动云检测. 信号

- 处理, 2019, 35(4): 556-562.
- An Jie, Ma Jinwen. Automatic cloud segmentation based on the fully convolutional neural networks. *Journal of Signal Processing*, 2019, 35(4): 556-562.
- [45] Chang Nibin, Bai Kaixu, Chen Chifarn. Smart information reconstruction via time-space-spectrum continuum for cloud removal in satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(5): 1898-1912.
- [46] Moosavi A, Attia A, Sandu A. A Machine Learning Approach to Adaptive Covariance Localization. Preprint at <https://arxiv.org/abs/1801.00548>, 2018: 1-24.
- [47] Cintra R, de Campos Velho H, Cocke S. Tracking the Model, Data Assimilation by Artificial Neural Network // 2016 International Joint Conference on Neural Networks (IJCNN), 2016: 403-410.
- [48] Lee Y J, Hall D, Stewart J, et al. Machine learning for targeted assimilation of satellite data. *Machine Learning and Knowledge Discovery in Databases*, 2018, 11053: 53-68.
- [49] Scher S. Toward data-driven weather and climate forecasting approximating a simple general circulation model with deep learning. *Geophys Res Lett*, 2018, 45(22): 12616-12622.
- [50] Brenowitz N D, Bretherton C S. Prognostic validation of a neural network unified physics parameterization. *Geophys Res Lett*, 2018, 45: 6289-6298.
- [51] Pan B, Hsu K, AghaKouchak A, et al. Improving precipitation estimation using convolutional neural network. *Water Resources Research*, 2019, 55(3): 2301-2321.
- [52] O'Gorman P A, Dwyer J G. Using machine learning to parameterize moist convection: Potential for modeling of climate, climate change, and extreme events. *Journal of Advances in Modeling Earth Systems*, 2018, 10(10): 2548-2563.
- [53] Rasp S, Pritchard M S, Gentile P. Deep learning to represent subgrid processes in climate models. *Proceedings of the National Academy of Sciences*, 2018, 115(39): 9684-9689.
- [54] Xu H, Zhang T, Luo Y, et al. Parameter calibration in global soil carbon models using surrogate-based optimization. *Geoscientific Model Development*, 2018, 11(7): 3027-3044.
- [55] Wu L, Zhang T, Qin Y, et al. An effective parameter optimization with radiation balance constraint in CAM5. *Geophys Res Lett*, 2020, 13: 41-53.
- [56] Burke A. Calibration of machine learning-based probabilistic hail predictions for operational forecasting. *Bull Amer Meteor Soc*, 2020, 35: 149-168.
- [57] Taillardat M. Calibrated ensemble forecasts using quantile regression forests and ensemble model output statistics. *Mon Wea Rev*, 2016, 144(6): 2375-2393.
- [58] Rasp S, Lerch S. Neural networks for postprocessing ensemble weather forecasts. *Mon Wea Rev*, 2018, 146(11): 3885-3900.
- [59] Scher S, Messori G. Predicting weather forecast uncertainty with machine learning. *Quart J Roy Meteor Soc*, 2018, 144(717): 2830-2841.
- [60] Sonderby C K, Espeholt L, Heek J, et al. MetNet: A Neural Weather Model for Precipitation Forecasting. Preprint at <https://arxiv.org/abs/2003.12140>, 2020: 1-17.
- [61] Ham Y G, Kim J H, Luo J J. Deep learning for multi-year ENSO forecasts. *Nature*, 2019, 573(7775): 568-572.
- [62] Zhou K, Zheng Y, Li B, et al. Forecasting different types of convective weather: a deep learning approach. *J Meteor Res*, 2019, 33(5): 797-809.
- [63] Kurth T, Treichler S, Romero J, et al. Exascale Deep Learning for Climate Analytics // Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis, 2018: 1-12.
- [64] Rojek K. Machine learning method for energy reduction by utilizing dynamic mixed precision on GPU-based supercomputers. *Concurrency and Computation: Practice and Experience*, 2019, 31(6): e4644. 1-e4644. 12.
- [65] Manandhar S, Dev S, Lee Y H, et al. A data-driven approach for accurate rainfall prediction. *IEEE Trans Geosci Remote Sens*, 2019, 57(11): 9323-9331.
- [66] Gagne D, Haupt S, Nychka D, et al. Interpretable deep learning for spatial analysis of severe hailstorms. *Mon Wea Rev*, 2019, 147(8): 2827-2845.
- [67] Karevan Z, Suykens J. Transductive LSTM for time-series prediction: An application to weather forecasting. *Neural Networks*, 2020, 125: 1-9.
- [68] Qiu M, Zhao P, Zhang K, et al. A Short-term Rainfall Prediction Model Using Multi-task Convolutional Neural Networks // 2017 IEEE International Conference on Data Mining, 2017: 395-404.
- [69] Yuan M, Ji X, Lu T, et al. A Novel Two-Factor Attention Encoder-Decoder Network through Combining Temporal and Prior Knowledge for Weather Forecasting // 2019 International Joint Conference on Neural Networks, 2019: 1-8.
- [70] Prasetya E P, Djamel E C. Rainfall Forecasting for the Natural Disasters Preparation Using Recurrent Neural Networks // 2019 International Conference on Electrical Engineering and Informatics, 2019: 52-57.
- [71] Tan C, Feng X, Long J, et al. FORECAST-CLSTM: A New Convolutional LSTM Network for Cloudage Nowcasting // 2018 IEEE Visual Communications and Image Processing, 2018: 1-4.

## Application of Artificial Intelligence Technology to Numerical Weather Prediction

Sun Jian<sup>1)</sup> Cao Zhuo<sup>2)</sup> Li Heng<sup>2)</sup> Qian Simeng<sup>3)</sup> Wang Xin<sup>2)</sup> Yan Limin<sup>2)</sup> Xue Wei<sup>2)3)4)5)</sup>

<sup>1)</sup> (National Meteorological Center, Beijing 100081)

<sup>2)</sup> (Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

<sup>3)</sup> (National Supercomputing Center in Wuxi, Wuxi 214072)

<sup>4)</sup> (Department of Earth System Science, Tsinghua University, Beijing 100084)

<sup>5)</sup> (Ministry of Education Key Laboratory for Earth System Modeling, Beijing 100084)

### Abstract

Numerical weather prediction technology plays an increasingly important role in improving accuracy and service level of modern weather forecast. With the development of observation system and higher resolution and complexity of the numerical weather prediction model, the products of numerical weather forecast have been greatly improved in quantity and quality, and can offer rich information at high spatial-temporal frequency. However, such a large amount of prediction data are not fully explored. Artificial intelligence has achieved great success in many fields, such as pattern recognition and natural language processing, which provides an opportunity for further improving numerical weather prediction. It's also employed in initialization, numerical model and production of weather forecast service, involving observation system, data assimilation, model integration, ensemble forecast and high-performance computing methods. Both the accuracy of forecast results and computational efficiency have been improved by using error correction, parameter estimation, local surrogate model and so on. In addition, some end-to-end neural network models also show the potential of pure data-driven weather forecast. These models use spatial-temporal observation data as input and directly output the prediction results in terms of deterministic results or probabilities. Some of them perform well in short-term severe convective weather, precipitation, and long-term climate forecast. Existing works employ various artificial intelligence technology methods, mainly including large-scale calculation of neural network, feature analysis, interpretability, and customized loss function. However, there are still some challenges, the potential of artificial intelligence needs to be further explored. Some issues should be carefully considered, including weak interpretability, uncertainty analysis and the coupling with conventional numerical models, and how to use physical knowledge to guide the design of artificial intelligence model is also worth addressing. To deal with these challenges, some promising suggestions are proposed. Bayesian network and causal network will help to establish more comprehensive and profound feature engineering. Using Bayesian inference to generate distribution characteristics of current meteorological states may be an alternative to efficient and effective uncertainty quantification. The development of some standard workflow and framework will contribute to the coupling of conventional numerical model and artificial intelligence module. Successful artificial intelligence applications in weather forecast require deep cooperation between meteorological experts and computer experts who focus on artificial intelligence and high-performance computing.

**Key words:** numerical weather prediction; artificial intelligence; machine learning; neural networks