

李妮娜,金荣花,宫宇,等,2026.“风清”人工智能天气预报模型业务应用性能评估[J].气象,52(5):513-526. Li N N, Jin R H, Gong Y, et al, 2026. Performance evaluation on operational application of the AI-based global short- and medium-range forecasting system—Fengqing Model[J]. Meteor Mon, 52(5):513-526 (in Chinese).

## “风清”人工智能天气预报模型业务应用性能评估\*

李妮娜<sup>1,2</sup> 金荣花<sup>1</sup> 宫宇<sup>1</sup> 代刊<sup>1</sup> 曹勇<sup>1,2</sup> 聂高臻<sup>1,2</sup>  
林建<sup>1</sup> 韦青<sup>1</sup> 吴海旭<sup>3</sup> 罗华坤<sup>3</sup> 龙明盛<sup>3</sup> 王建民<sup>3</sup>

1 国家气象中心,北京 100081

2 河北省气象人工智能重点实验室,雄安气象人工智能创新研究院,雄安新区 070001

3 清华大学,北京 100084

**提 要:** 2024 年,中国气象局与清华大学联合研发了基于“AI+物理”的“风清”人工智能天气预报模型(简称风清模型)。该模型采用多尺度隐空间投影架构与能量守恒损失函数等设计,实现了对全球短中期天气的智能预报,并已投入业务化运行。本研究基于预报准确率、偏差空间分布等指标,全面评估了 2024 年该模型在中国及周边区域的预报能力;并聚焦台风和降水两类典型天气过程,重点检验了其在灾害性天气业务预报中的性能表现。结果表明,风清模型 500 hPa 位势高度场的有效预报时效超过 10 d,地面 2 m 气温及高空 850 hPa 温度预报的均方根误差较欧洲中期天气预报中心 ECMWF-IFS 模式显著偏小(最大改进幅度达 37.66%);从典型天气过程来看,风清模型对台风路径的预报略优于 ECMWF-IFS 模式,但台风强度预报偏弱;同时,该模型对暴雨有较好的预报能力,对台风降水及梅雨雨带的落区预报误差较小,暴雨预报 TS 评分在 73~168 h 时效下较 ECMWF-IFS 模式提升了 43.53%。总体而言,风清模型在业务预报中展现出较高潜力,但中长期时效下的活跃度与台风强度预报有待进一步提升。

**关键词:** 风清模型,检验评估,灾害性天气

**中图分类号:** P456

**文献标志码:** A

**DOI:** 10.7519/j.issn.1000-0526.2026.030501

## Performance Evaluation on Operational Application of the AI-Based Global Short- and Medium-Range Forecasting System—Fengqing Model

LI Nina<sup>1,2</sup> JIN Ronghua<sup>1</sup> GONG Yu<sup>1</sup> DAI Kan<sup>1</sup> CAO Yong<sup>1,2</sup> NIE Gaozhen<sup>1,2</sup>  
LIN Jian<sup>1</sup> WEI Qing<sup>1</sup> WU Haixu<sup>3</sup> LUO Huakun<sup>3</sup> LONG Mingsheng<sup>3</sup> WANG Jianmin<sup>3</sup>

1 National Meteorological Centre, Beijing 100081

2 Hebei Key Laboratory of Meteorological Artificial Intelligence, Xiong'an Institute of Meteorological Artificial Intelligence, Xiong'an New Area 070001

3 Tsinghua University, Beijing 100084

**Abstract:** In 2024, China Meteorological Administration (CMA), in collaboration with Tsinghua University, developed the forecasting model following an innovative “AI-Physics” hybrid approach—Fengqing Model. With the designs such as a multi-scale latent space projection architecture and an energy-conservation loss function, the model has been equipped with global short- and medium-range weather forecasting capabilities and has been applied in forecasting operations. This paper comprehensively evaluates the forecasting

\* 新疆维吾尔自治区重点研发专项(2022B03027-2)、国家自然科学基金项目(42405019)、中国气象局创新发展专项(CXFZ2025J086)及青年创新团队(CMA2023QN05)共同资助

2025 年 4 月 30 日收稿; 2026 年 3 月 15 日收修定稿

第一作者:李妮娜,主要从事精细化检验评估研究. E-mail: linn@cma.gov.cn

通讯作者:宫宇,主要从事天气学诊断分析研究. E-mail: heygy123@126.com

ability of Fengqing Model in China and the surrounding areas in 2024 from several metrics such as forecasting accuracy and bias distribution. Two kinds of typical synoptic processes, i. e., typhoon and rainstorm, are also focused on deeply exploring the model's performance in the forecasts of disastrous weather. The results show that the 500 hPa geopotential height forecasts by Fengqing Model maintain a predictive skill beyond 10 days. The root mean square errors of 2 m temperature and 850 hPa temperature are significantly lower than those from the European Centre for Medium-Range Weather Forecasts (ECMWF-IFS), having a maximum improvement of 37.66%. In terms of typical weather processes, Fengqing Model demonstrates its superior performance in typhoon track forecasting to ECMWF-IFS, but it underestimates the typhoon intensity. In addition, Fengqing Model has a good forecasting ability for rainstorm, with smaller forecast errors for the locations of typhoon rain and Meiyu front rainfall belt. The TS score of rainstorm forecasts in the medium-range (73–168 h lead time) is improved by 43.53% compared to that of ECMWF-IFS forecasts. Overall, the Fengqing Model presents considerable potential in operational forecasting, although further improvements are needed in activity level and typhoon intensity prediction at medium- and long-range lead times.

**Key words:** Fengqing Model, verification and evaluation, disastrous weather

## 引 言

近年来,人工智能技术的突破推动了气象预报领域技术革新(Schultz et al, 2021; Ebert-Uphoff and Hilburn, 2023; Nguyen et al, 2023; 马雷鸣, 2020; 陈鹤等, 2022; 黄小猛等, 2024; 周康辉等, 2025), 基于深度学习构建的数据驱动天气预报模型成为研究热点(Rasp et al, 2020; 杨绚等, 2022; 盛杰等, 2025)。国际上, 以谷歌的 GraphCast (Lam et al, 2023)、英伟达的 FourCastNet (Pathak et al, 2022) 以及华为的盘古气象模型(Bi et al, 2023) 为代表的人工智能天气预报模型, 已展现出对传统数值模式在计算效率与部分要素预报精度上的显著优势(Chen et al, 2023; Lam et al, 2023)。与此同时, 我国气象部门积极布局智能化转型, 2024 年中国气象局联合清华大学发布首个天气预报模型——人工智能全球中短期预报系统“风清”(以下简称风清模型)。区别于国内外主流研究多采用的视觉变换器网络(Bojesomo et al, 2023; Nguyen et al, 2023)、图神经网络(Keisler, 2022; Lam et al, 2023)、神经算子网络、扩散生成模型等技术路线, 风清模型从“AI+物理”的路线出发, 创新地将大气物理深度融入深度学习模型设计, 构建了具有全新架构且具备强物理表征能力的模型结构。该模型已于 2024 年 9 月实现业务化运行。

开展预报检验评估是深入认识人工智能模型性能并推动其发展的重要环节(Rasp et al, 2020; Rasp

and Thuerey, 2021; Radford et al, 2025)。然而, 当前的检验多聚焦于模型的全球尺度性能, 且主要基于统计检验指标(距平相关系数和均方根误差等)进行(Bi et al, 2023)。考虑到人工智能模型与传统基于物理的数值模式存在本质差异, 亟需从天气学的视角出发, 针对性地评估模型对典型天气过程的预报能力(Ebert-Uphoff and Hilburn, 2023; Radford et al, 2025)。此外, 关于区域精细化预报能力的系统性检验尚存在空白。特别是针对风清模型, 其在我国及周边地区对极端天气的捕捉能力、误差分布特征及其对预报员决策的辅助能力尚未明晰。为此, 本文从预报精度、偏差分布和典型天气过程发展演变预报性能等方面入手, 全面评估风清模型在中国及其周边区域的业务化表现, 相关结论可为模型应用策略的优化以及天气预报业务的实际应用提供科学参考。

## 1 数据和方法

### 1.1 风清模型技术特点

风清模型遵循人工智能与物理约束相结合的总体思路, 基于对大气多尺度动力过程的系统认知, 自主构建了一套融合物理可解释性的深度学习天气预报框架。模型创新地提出了多尺度隐空间投影的总体架构(Wu et al, 2023), 将天气系统的多尺度时空耦合过程转化为隐空间中的状态转移问题, 从而在保留动力学一致性的同时有效降低模型训练复杂

度。该设计实现了从数据驱动到物理约束的统一，使模型能够在隐空间中学习并归纳大气活动的内在规律，增强了预报结果的物理可解释性。

在损失函数构建上，模型结合灾害性天气形成过程中的多尺度相互作用机制及西风带环流基本特征，构建了多尺度特征提取网格，并引入侧重同纬度上下游关系的注意力机制。针对数据驱动模型中常见的“过平滑”及能量耗散问题，模型引入能量守恒隐式表征模型进行约束，从粗网格分析场中提取气象变量的守恒性质，以改善长时效预报的活跃及对极端天气过程的再现能力。

在训练策略方面，模型提出了可扩展的多时效联合优化策略，通过分离计算不同预报时效的参数梯度并进行统一优化，减少了迭代误差积累，并兼顾短期和中期预报效果。该策略可有效延长可训练预报时效，并提高模型在不同时间尺度上的稳定性与一致性。

在降水预报部分，风清模型针对 6 h 累计降水的生成机制与统计特征，自主建立了协变量驱动的降水预测框架，将降水自身的历史演变特征与大气背景场的关系相结合进行学习。为改善强降水样本稀疏性问题，采用增加强降水预报训练权重及中国区域主汛期时段样本重采样策略，进而提升对强降水的预报效果。

另外，在灾害性天气的预报优化方面，针对诸如台风、暴雨雪等灾害性天气以及气旋、锋面等天气系统发生的区域识别优化，根据对其产生的天气学条件分析，发现其发生均与强风的出现有着密切关联（梅婵娟等，2025；邱孙俊杰和张芳华，2025；孙一昕和张楠，2025）。基于这一认识，通过对大风出现区域加权优化，能够间接客观识别灾害性天气发生位置，从而实现针对性的预报优化。

模型的训练基于 1979—2017 年（共 39 年）ERA5 再分析资料，依托中国气象局国家气象信息中心气象人工智能基础支撑技术平台完成。所得模型包含约 30 亿个参数，能够在约 3 min 内生成未来 15 d、25 km 分辨率、逐 6 h 的全球气象预报产品。

预报产品涵盖地面要素（海平面气压、2 m 气温、10 m 经向和纬向风）和高空要素（1000~50 hPa 共 13 个气压层的位势高度、气温、比湿、经向和纬向风）。该模型（CMA-AIM-GFS-Fengqing-V1.0 版本）已成功通过中国气象局业务化准入流程，面向全国气象业务部门实时应用。

## 1.2 资料

本文针对风清模型的预报结果进行检验，同时将其预报结果与欧洲中期天气预报中心 ECMWF-IFS 全球数值模式的 C1D 版本预报结果（以下简称 ECMWF 模式）进行对比评估。时段为 2024 年 1 月 1 日至 12 月 31 日，检验对象包括逐日 08 时（全文均为北京时间）和 20 时起报的 6~240 h 逐 6 h 预报产品，要素包括 500 hPa 位势高度场、700 hPa 湿度场、850 hPa 温度场和风场、2 m 气温、10 m 风场、海平面气压以及降水。其中，风清模型的空间分辨率为 25 km，ECMWF 模式地面预报产品的空间分辨率为 12.5 km，气压层预报产品空间分辨率为 25 km。

本文所用的风清模型的产品信息如表 1 所示。对于除降水外的气象要素，采用中国气象局业务准入的 CMA-AIM-GFS-Fengqing-V1.0 版本，严格遵循实际业务应用场景（以 ECMWF 模式零场和 6 h 预报场作为输入），对 2024 年开展逐 12 h 起报的 0~360 h 时效预报产品回算；对于降水预报，则基于在中央气象台内部测试应用的 CMA-AIM-GFS-Fengqing-beta 版本开展回算。考虑到风清模型于 2024 年 9 月正式业务化运行，为获取完整一年的预报样本以支撑系统性检验，按上述实际业务应用场景对 2024 年进行了回算。该回算结果在数据及算法框架上与业务预报完全一致，对其检验可等效反映模型在实际业务应用中的表现。

本文在开展格点检验时采用 ERA5 再分析资料（Hersbach et al., 2020）作为真值，相关要素包括 200 hPa 风场、500 hPa 位势高度场、700 hPa 湿度场、850 hPa 温度场和风场、海平面气压场、2 m 气温、10 m 风场。台风检验采用中央气象台台风路径

表 1 研究采用的风清模型版本及产品信息

Table 1 Version and product information of the Fengqing Model used in this article

| 版本                        | 产品要素   | 起报时间(BT)      | 预报时效    | 时间分辨率 | 空间分辨率 |
|---------------------------|--|---------------|---------|-------|-------|
| CMA-AIM-GFS-Fengqing-V1.0 | 200 hPa 风场、500 hPa 位势高度场、700 hPa 湿度场、850 hPa 温度场和风场、2 m 气温、10 m 风场、海平面气压 | 每日 08 时和 20 时 | 6~240 h | 6 h   | 25 km |
| CMA-AIM-GFS-Fengqing-beta | 降水量  |               |         |       |       |

和强度实时业务资料作为真值;降水检验采用中国气象局国家级气象台站逐小时降水观测数据,为了与逐 6 h 预报进行匹配,将台站观测的连续 6 个逐小时降水数据累加得到逐 6 h 累计降水,若时段内台站降水存在缺测,则该 6 h 累计降水记为缺测。

文中地图基于审图号为 GS(2019)3082 号的标准地图制作,底图无修改。

### 1.3 检验方法

针对关键要素的总体预报检验,聚焦我国及周边地区,区域范围为  $0^{\circ}\sim 60^{\circ}\text{N}$ 、 $70^{\circ}\sim 140^{\circ}\text{E}$ 。这部分检验主要在格点上进行。为了保持水平分辨率统一,研究将 ECMWF 模式的结果双线性插值至与风清模型一致的分辨率。检验指标包括距平相关系数 (ACC)、预报活跃度 (Forecast Activity)、均方根误差 (RMSE)。其中,预报活跃度是用于评估天气现象可变性的指标,能够直观反映预报相对于实况的“活跃”或“动态”程度;高活跃度意味着模型预报的天气过程频繁且变化较为剧烈;反之,则表示预报的天气相对平稳、变化较少且更接近气候平均态 (Bouallégue et al, 2024)。

相关指标计算公式如下:

$$\text{ACC} = \frac{\sum_{j=1}^{N_y} \sum_{i=1}^{N_x} \omega_j (f_{i,j} - c_{i,j})(o_{i,j} - c_{i,j})}{\sqrt{\left[ \sum_{j=1}^{N_y} \sum_{i=1}^{N_x} \omega_j (f_{i,j} - c_{i,j})^2 \right] \left[ \sum_{j=1}^{N_y} \sum_{i=1}^{N_x} \omega_j (o_{i,j} - c_{i,j})^2 \right]}}$$

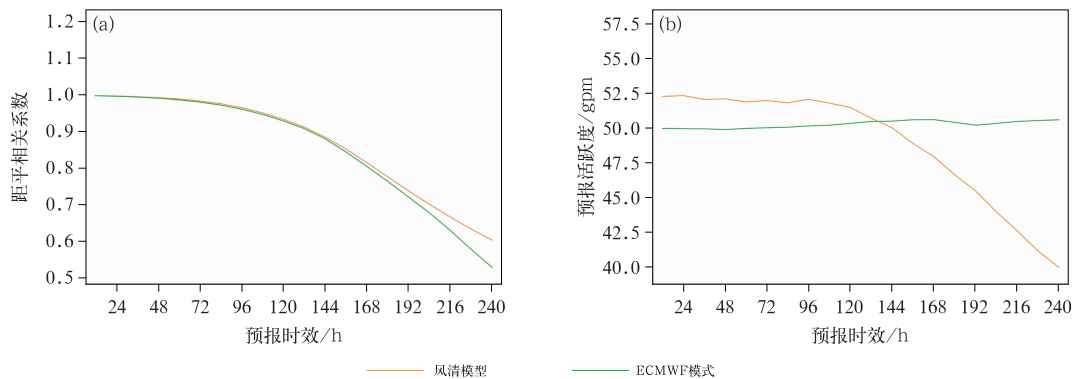


图 1 风清模型和 ECMWF 模式对 2024 年中国及周边区域 500 hPa 位势高度预报的 (a) 距平相关系数和 (b) 预报活跃度随预报时效的变化

Fig. 1 Variations of (a) anomaly correlation coefficient (ACC) and (b) forecast activity of the 500 hPa geopotential height forecasts by the Fengqing and ECMWF models for China and its surrounding region in 2024 with the forecast lead time

$$\text{Forecast Activity} = \sqrt{\frac{\sum_{j=1}^{N_y} \sum_{i=1}^{N_x} \omega_j (f_{i,j} - c_{i,j})^2}{\sum_{j=1}^{N_y} \sum_{i=1}^{N_x} \omega_j}}$$

式中:  $f_{i,j}$ 、 $o_{i,j}$ 、 $c_{i,j}$  分别为格点  $i, j$  处的预报、观测、气候态均值,  $N_x$  和  $N_y$  分别是经向 ( $i$ ) 和纬向 ( $j$ ) 的总格点数。  $\omega_j$  为第  $j$  个纬度的加权系数, 即  $\omega_j = \cos\varphi_j$ 。本文采用的气候态均值是基于 1989—2008 年的 ERA-40 资料, 利用 61 d 的窗口滑动平均得到。

对典型天气过程的预报检验, 主要采用双线性插值方法将不同要素预报产品插值到全国 2401 个国家级气象台站后开展。检验指标包括 RMSE、TS 评分、Bias 评分等。降水预报的二分类检验基于设定阈值 (如  $\geq 0.1 \text{ mm}$ ) 划分事件, 采用 TS 评分和 Bias 评分两个指标。其中: TS 评分越接近 1, 表明预报命中率越高、预报性能越优; Bias 评分越接近 1 预报越合理, 高于 1 表示预报过度, 空报次数多于漏报次数, 低于 1 表示预报不足, 漏报次数多于空报次数 (刘会军等, 2024)。

## 2 关键要素的总体预报检验

图 1 给出中国及周边地区 500 hPa 位势高度预报的 ACC 及预报活跃度。500 hPa 位势高度是评估天气预报技巧的重要变量之一, 通常将距平相关系数超过 0.6 的预报时效作为有效预报时效 (沈学顺等, 2021)。从图 1a 可以看出, 风清模型预报的

500 hPa 位势高度 ACC 优于数值模式 ECMWF 预报结果,尤其是在长期时效下优势突出。ECMWF 模式预报结果在 216 h 和 228 h 时效下的 ACC 分别为 0.630 和 0.579,即其有效预报天数超过 9 d 但不足 9.5 d;对于风清模型,240 h 时效下 ACC 为 0.604,即其预报的 500 hPa 位势高度在中国及周边区域最长有效预报时效超过 10 d。

Keisler(2022)研究表明,模型的训练会导致过于平滑的预报结果,从而使得 ACC 评分较高、RMSE 较小。因此在评估 ACC、RMSE 等指标时,有必要对预报结果的活跃水平进行比较(Bouallègue et al,2024)。由图 1b 可见,ECMWF 模式 500 hPa 位势高度的预报活跃度随着预报时效增加稳定维持,而风清模型预报活跃度在前 120 h 时效稳定维持,随后则随预报时效的增加明显下降。也就是说,在较长的预报时效下,风清模型对剧烈变化的天气预报能力较差。

图 2 给出高空以及地面关键要素预报 RMSE 随时效的变化。从图中可以看出,风清模型相较于 ECMWF 模式对不同要素的预报均有所改善,其中 500 hPa 位势高度和海平面气压预报结果的改善相对较小,尤其是 500 hPa 位势高度,在 96~156 h 预报时效下的风清模型的 RMSE 略高于 ECMWF 模式结果。对于其他要素的预报,风清模型预报效果改进幅度在 5.20%~37.66%。其中,2 m 气温的预报改进效果最为突出, RMSE 的最大降幅达 37.66%,即便是最小降幅也达到了 19.12%。

考虑到实际业务中通常关注 72 h 内预报效果,图 3 给出不同要素 0~72 h 预报 RMSE 空间分布。对于 200 hPa 风场来说(图 3a),风清模型对 20°N 以南地区以及我国东部上空的预报改善较为明显, RMSE 较 ECMWF 模式偏小 1~2 m·s<sup>-1</sup>。对于 500 hPa 位势高度场来说(图 3b),风清模型主要改善了中低纬地区的预报误差,尤其是对洋面上空的位势高度, RMSE 减小超过 3.2 gpm;在 25°N 以北地区,风清模型预报误差略高于 ECMWF 模式,尤其在我国青藏高原以及东北地区, RMSE 偏高 0.82 gpm(图略)。风清模型对 700 hPa 比湿的预报较 ECMWF 模式改进明显,在中低纬地区 RMSE 减小超过 0.4 g·kg<sup>-1</sup>,对我国东部地区比湿预报的 RMSE 也能减小 0.3 g·kg<sup>-1</sup>(图 3c)。从 850 hPa 温度预报 RMSE

空间分布可以看出,第二地形阶梯(海拔在 500~3000 m)上的温度预报误差较大,风清模型可减小该地区温度预报的预报误差, RMSE 在 1.5℃左右(图 3d);值得注意的是,风清模型对长江流域的预报改进更为突出,850 hPa 温度预报 RMSE 减小约 0.5℃。对于海平面气压来说(图 3f),ECMWF 模式在高海拔地区存在较大的预报误差,风清模型预报对该地区预报改进最为明显,尤其是青藏高原周边的 RMSE 可减小 200 Pa 以上。与海平面气压预报效果类似,风清模型预报的 2 m 气温在青藏高原周边及天山山脉处的 RMSE 较 ECMWF 模式偏小 1℃以上(图 3g)。从图 3e 和图 3h 可以看出,ECMWF 模式预报 850 hPa 风速误差最大区域位于日本海附近,其预报的 10 m 风速误差大值位于沿海、天山山脉地区,风清模型预报可有效改善上述地区的预报误差。总体而言,风清模型预报误差的空间分布较 ECMWF 模式偏小,对洋面、海陆交界以及高海拔地区的预报误差改善较为明显。

通过对 2024 年关键要素的总体预报检验可知,风清模型对我国及周边地区环流形势的预报效果较好,预报误差较 ECMWF 模式有所减小,500 hPa 位势高度场最长有效预报天数超 10 d。但相较于全球尺度预报来说,在我国及周边地区的预报仍存在一定偏差,未来仍需关注 500 hPa 位势高度场的预报,一方面应注意提高中长期预报的活跃度,另一方面应注意中高纬地区的预报改善。对于地面要素 2 m 气温、10 m 风速来说,风清模型预报改进明显,尤其对前者的改进幅度较为突出,针对 ECMWF 模式预报误差较大的区域,可参考风清模型预报结果。

### 3 台风和降水预报检验

考虑到仅评估总体预报水平不足以了解模型在实际业务预报中的应用价值,本节将重点针对 2024 年发生的典型天气过程(主要关注台风、暴雨)进行细致检验,旨在揭示预报中的具体偏差特征,为后续优化改进模型提供依据。

#### 3.1 台风

##### 3.1.1 2024 年整体预报检验

基于 2024 年全年 26 个编号台风的路径预报进

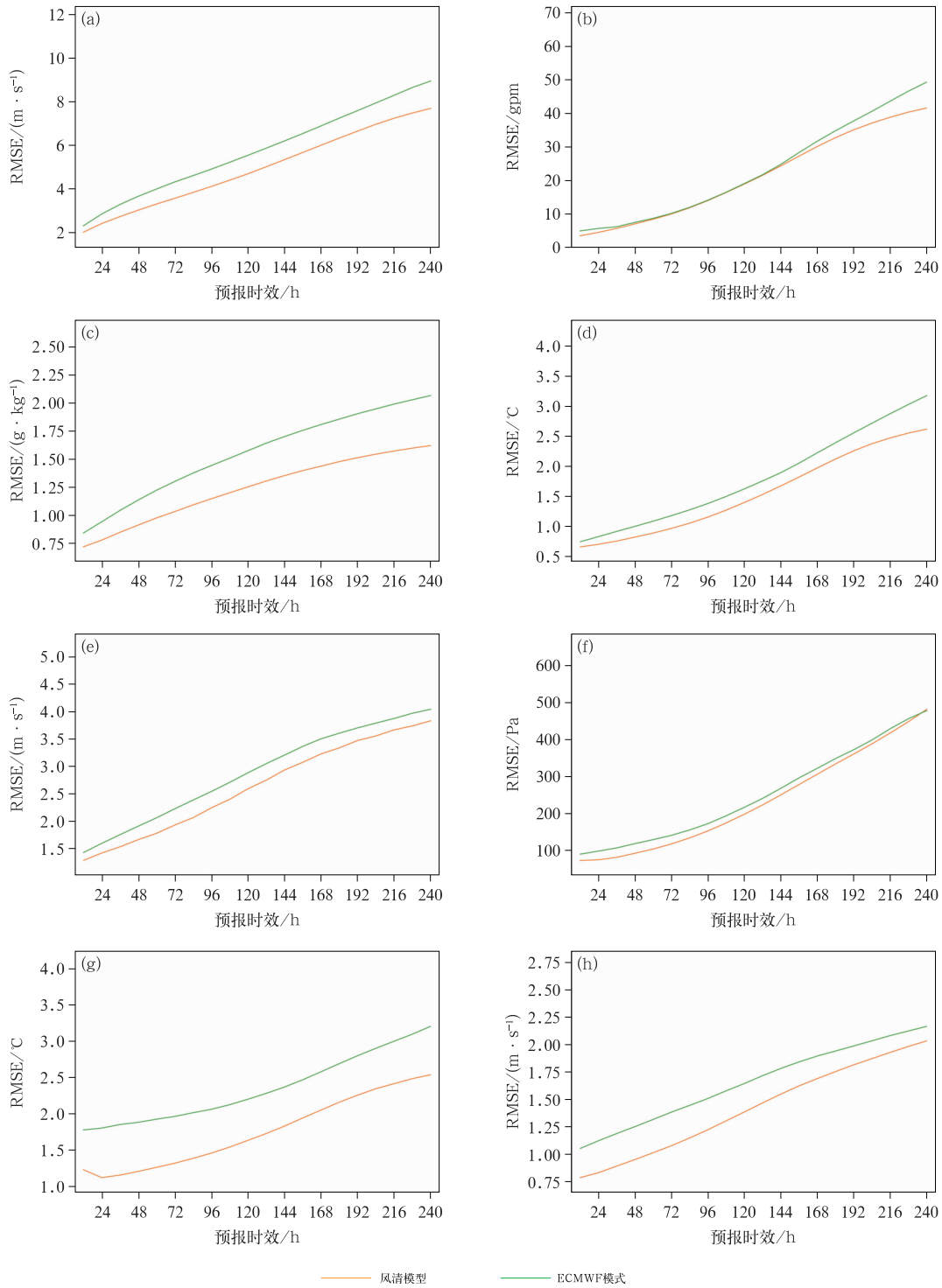


图 2 风清模型和 ECMWF 模式对 2024 年中国及周边区域

(a)200 hPa 风速,(b)500 hPa 位势高度,(c)700 hPa 比湿,(d)850 hPa 温度,

(e)850 hPa 风速,(f)海平面气压,(g)2 m 气温和(h)10 m 风速预报 RMSE 随预报时效的变化

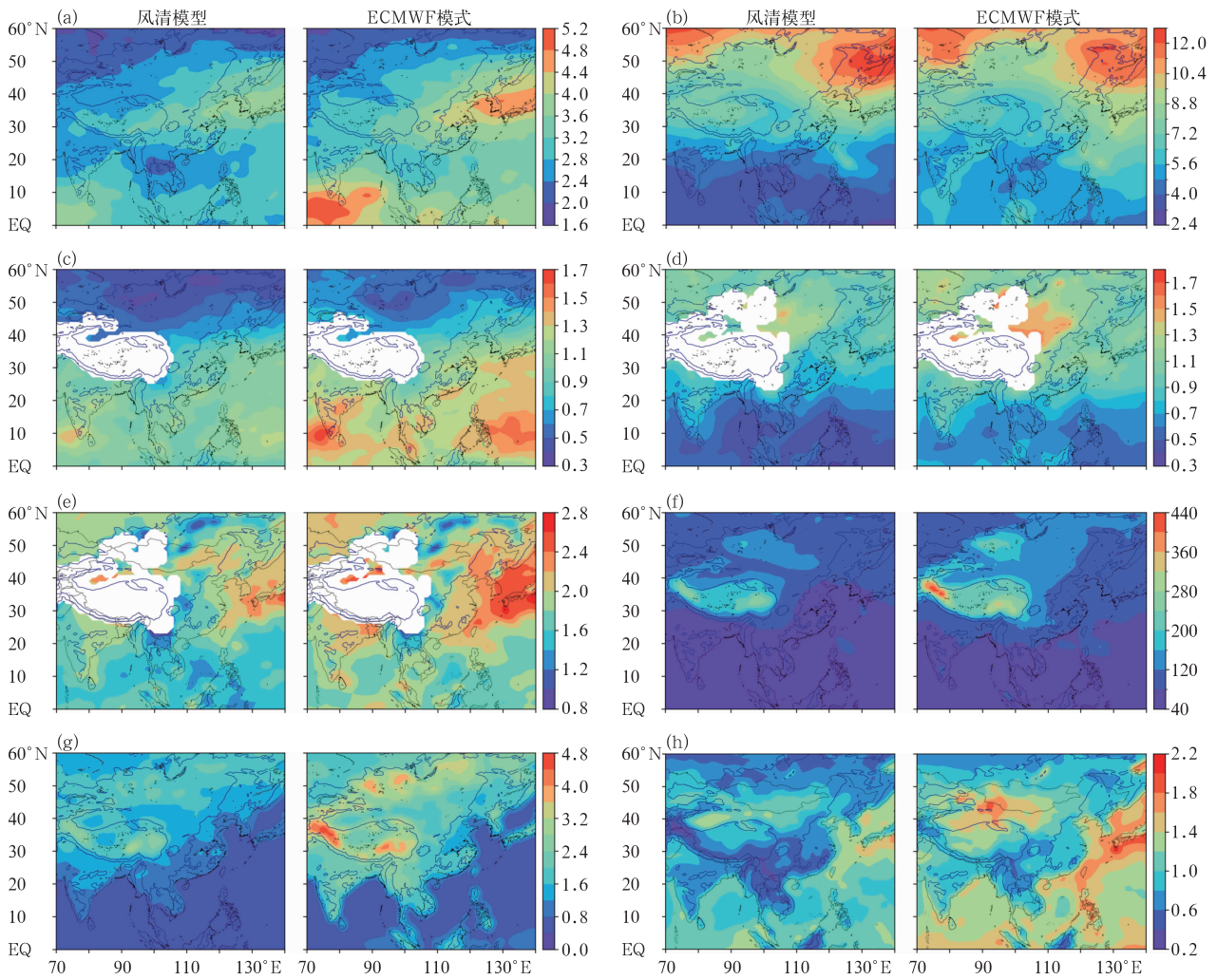
Fig. 2 Variations with the forecast lead time of root mean square error (RMSE) of (a) 200 hPa

wind speed, (b) 500 hPa geopotential height, (c) 700 hPa specific humidity, (d) 850 hPa temperature,

(e) 850 hPa wind speed, (f) sea-surface pressure, (g) 2 m temperature and (h) 10 m wind speed by Fengqing and ECMWF

models for China and its surrounding region in 2024

models for China and its surrounding region in 2024



注：蓝色等值线标注 500 m 和 3000 m 地形高度，空白区域为地形高度高于对应等压面的掩蔽区域。

图 3 风清模型和 ECMWF 模式对 2024 年中国及周边区域(a)200 hPa 风场(单位： $\text{m} \cdot \text{s}^{-1}$ )，  
(b)500 hPa 位势高度(单位：gpm)，(c)700 hPa 比湿(单位： $\text{g} \cdot \text{kg}^{-1}$ )，  
(d)850 hPa 温度(单位： $^{\circ}\text{C}$ )，(e)850 hPa 风速(单位： $\text{m} \cdot \text{s}^{-1}$ )，(f)海平面气压(单位：Pa)，  
(g)2 m 气温(单位： $^{\circ}\text{C}$ )和(h)10 m 风速(单位： $\text{m} \cdot \text{s}^{-1}$ )72 h 预报的 RMSE 空间分布  
Fig. 3 Spatial distribution of RMSE of 72 h forecasts of (a) 200 hPa wind speed (unit:  $\text{m} \cdot \text{s}^{-1}$ ),  
(b) 500 hPa geopotential height (unit: gpm), (c) 700 hPa specific humidity (unit:  $\text{g} \cdot \text{kg}^{-1}$ ),  
(d) 850 hPa temperature (unit:  $^{\circ}\text{C}$ ), (e) 850 hPa wind speed (unit:  $\text{m} \cdot \text{s}^{-1}$ ), (f) sea-surface  
pressure (unit: Pa), (g) 2 m temperature (unit:  $^{\circ}\text{C}$ ), and (h) 10 m wind speed (unit:  $\text{m} \cdot \text{s}^{-1}$ )  
by Fengqing and ECMWF models for China and its surrounding region in 2024

行评估,对比 ECMWF 模式与风清模型预报情况(检验样本量相同)。从图 4a 能够看出,在 108 h 以内,台风路径误差随预报时效缓慢增长,在此期间风清模型路径预报效果略优于 ECMWF 模式,预报误差在 53~283 km。特别在 72 h 时效时,风清模型的预报效果相较 ECMWF 模式优势最为明显,误差减小约 14%。在 108~240 h 时效,路径预报误差增

长速度较快,此时风清模型预报效果略差于 ECMWF 模式。通常而言,对于台风路径的预报,人工智能模型相较于数值模式在不同程度上有所提升,而对于台风强度的预报则普遍较数值模式预报结果偏弱(Bouallègue et al,2024)。就风清模型来说,虽然其台风强度预报同样较 ECMWF 模式偏弱,在各预报时效下均呈现负误差,24 h 预报的强度误差为

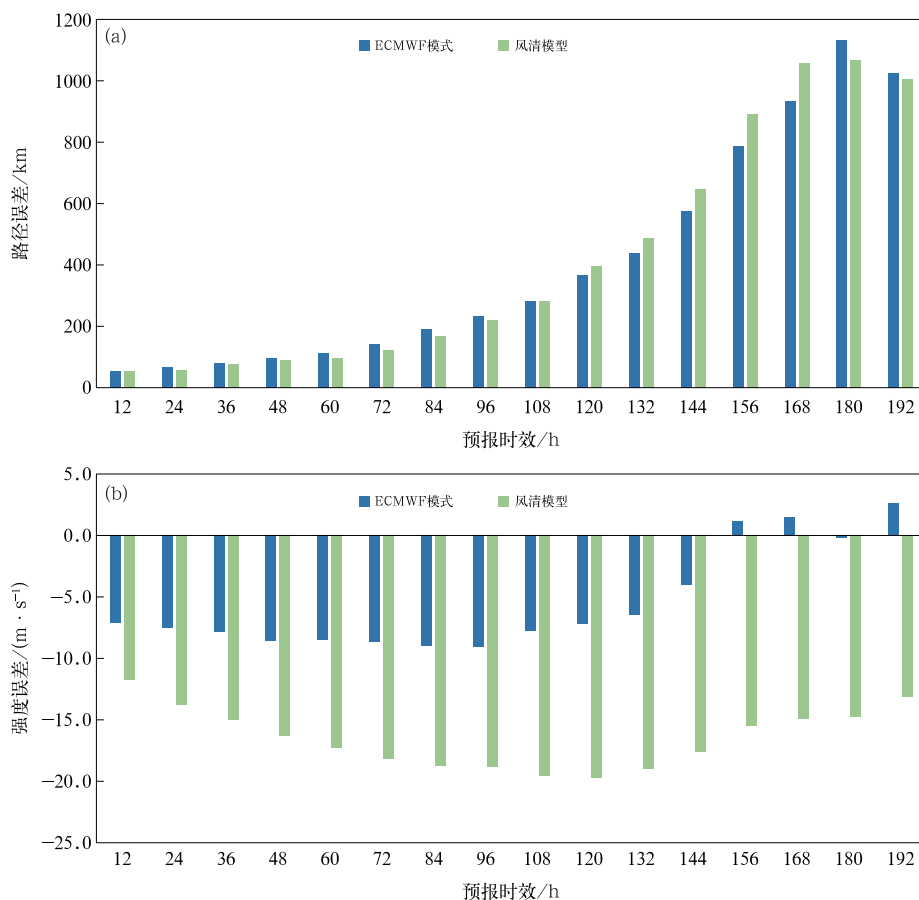


图4 风清模型和 ECMWF 模式对 2024 年台风的(a)路径和(b)强度预报误差  
Fig. 4 Forecast errors of Fengqing and ECMWF models for (a) typhoon track and (b) typhoon intensity in 2024

$-11.8 \text{ m} \cdot \text{s}^{-1}$ , 但该预报结果仍略优于伏羲等人工智能模型预报结果(Niu, 2025)。另外, 从台风中心最低海平面气压场的预报来看, 风清模型 0~72 h 时效内预报误差仅在 25 hPa 左右(图略), 而风乌、盘古、伏羲等人工智能模型的误差接近 30 hPa(Liu et al, 2024)。风清模型通过对大风出现区域进行加权优化, 在一定程度上减缓了强度预报误差, 但未来模型优化仍需关注更多与台风强度有关的物理条件以提升其预报水平。

### 3.1.2 “格美”台风个例预报分析

台风“格美”起编前, 风清模型和 ECMWF 模式即较好地预报出了台风未来移动方向, 且对于台风在菲律宾以东洋面的路径转向也有所反映。虽然二者均未能预报出登陆台湾, 但是都提前 5 d 以上预报出了其将在东南沿海登陆。值得注意的是: (1) 风清模型预报路径在海上阶段更接近实况, (2) 此时风清模型已预报出台风登陆后西行的路径(图 5a)。

21 日 20 时, “格美”为强热带风暴级, 此时风清

模型和 ECMWF 模式预报路径、登陆点(台湾和福建两次登陆)均与实况已相当接近, 登陆台湾之前的阶段, 风清模型的预报路径和实况基本重合。主要分歧在于登陆之后, 风清模型仍预报台风向西移动, 影响湖南等地, 实况是“格美”在湖南引发灾害性降水; 而 ECMWF 模式预报台风登陆后经安徽北上, 进入山东(图 5b)。

### 3.1.3 “摩羯”台风个例预报分析

台风“摩羯”是 2024 年最强的台风之一, 给海南、广东带来严重影响。9 月 1 日 20 时, 在“摩羯”正式起编前, 该系统仍为位于菲律宾以东洋面的热带低压。此时风清模型和 ECMWF 模式均报出了台风先向北、后转向西移动并影响我国华南的趋势。相比来看, 风清模型预报与实况更加接近。在进入南海后, ECMWF 模式预报台风纬度有所南落, 与实际偏差较大。分歧更大的是后期路径, 接近海南岛时 ECMWF 模式预报台风路径北抬, 最终进入广西内陆地区, 与实况不符(图 6a)。

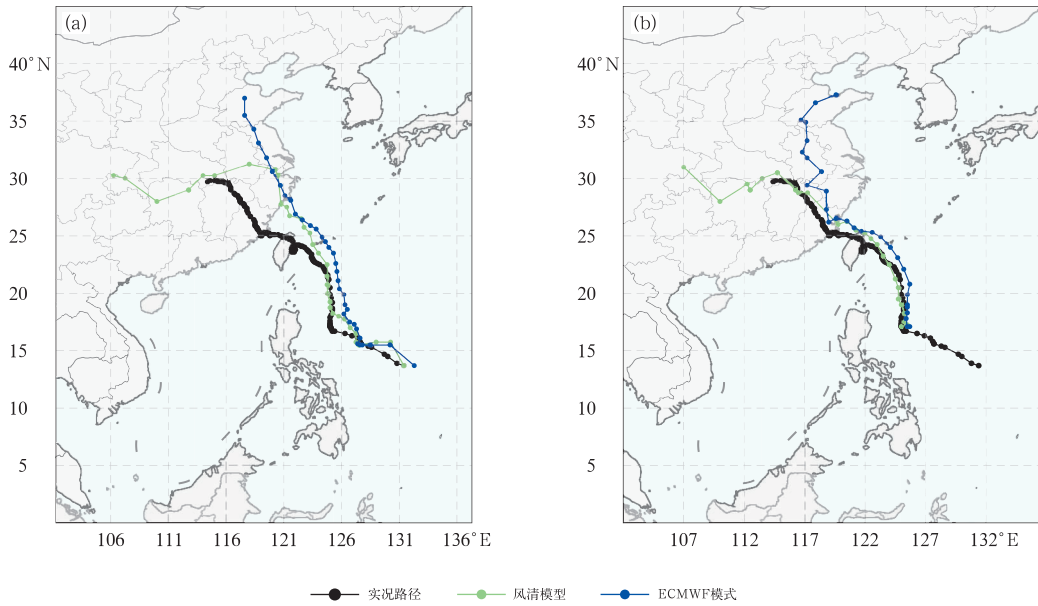


图 5 风清模型和 ECMWF 模式 2024 年 7 月(a)19 日 20 时起报、  
(b)21 日 20 时起报的台风“格美”路径预报与实况的比较

Fig. 5 Comparison of the observed track and predicted track by Fengqing and ECMWF models of Typhoon Gaemi initialized at (a) 20:00 BT 19 and (b) 20:00 BT 21 July 2024

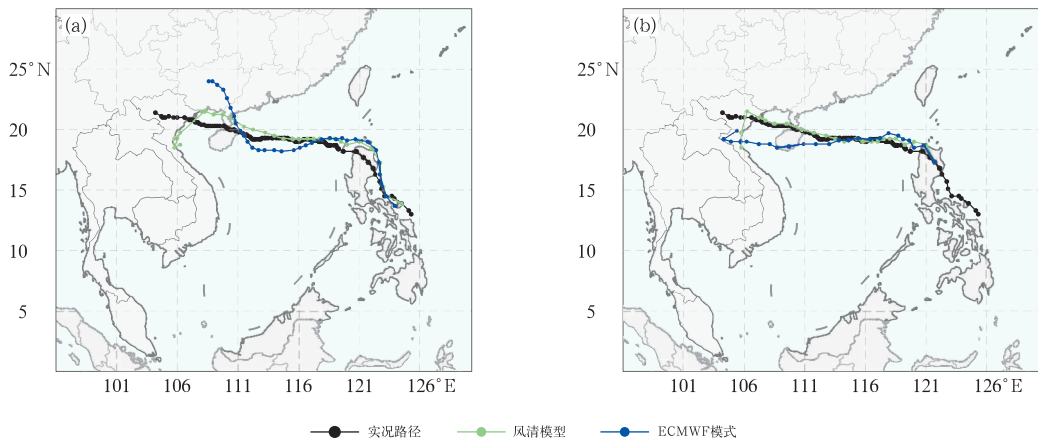


图 6 风清模型和 ECMWF 模式 2024 年 9 月(a)1 日 20 时起报、  
(b)2 日 20 时起报的超强台风“摩羯”路径预报与实况的比较

Fig. 6 Comparison of the observed track and predicted track by Fengqing and ECMWF models of super Typhoon Yagi initialized at (a) 20:00 BT 1 and (b) 20:00 BT 2 September 2024

24 h 后(9 月 2 日 20 时)的预报,风清模型的台风路径与实况吻合度相当好,准确预报了在海南岛和广东雷州半岛的两次登陆。ECMWF 模式预报路径相比之前明显地向南调整,与实况仍然有较大偏差,不仅未能预报出在广东的登陆,且在海南岛的登陆位置与实况也有不小的差距(图 6b)。此后随着台风逐渐西移进入南海,风清模型和 ECMWF 模式预报趋于一致,结果均较好(图略)。

综上,不论是西北行登陆并深入内陆的“格美”,

还是南海西行的超强台风“摩羯”,风清模型相比 ECMWF 模式均有更好的预报表现。统计结果也表明,风清模型在 120 h 时效以内表现突出,路径误差整体小于 ECMWF 模式;然而,120 h 后,其预报误差反而超过后者,表明风清模型在中长期预报中的准确性仍存在不足。这种性能差异可能与其在台风强度预报上的系统性偏弱、中长期预报场表现出过度平滑等问题有关。已有研究指出,预报场平滑化是人工智能天气预报模型的普遍特征,容易削弱

模型对中小尺度系统的刻画能力,进而影响其对引导气流强度和路径突变等关键动力过程的响应(Liu et al,2024;Chen et al,2025)。此外,受限于物理机制表达能力和训练样本的代表性问题,当前人工智能模型在应对双台风相互作用、路径急转弯等复杂过程时能力仍显不足(Xu et al,2025)。这些结果表明,尽管人工智能模型在台风短期预报中展现出一定优势,但提升其对复杂动力过程的表征能力,仍需通过引入物理约束、扩展数据多样性以及优化模型训练策略等方式加以强化(Olivetti and Messori, 2024)。

### 3.2 降水

#### 3.2.1 2024 年整体预报检验

图 7 首先给出 2024 年降水预报的整体情况。从晴雨预报准确率来看(图 7a),大部分时效下风清模型和 ECMWF 模式预报准确率均超过 0.5,风清

模型略低于 ECMWF 模式。但对于大雨和暴雨的预报,风清模型 TS 评分明显更高,0~72 h 时效下暴雨预报 TS 评分综合来看超过 0.21,较 ECMWF 模式提升了 27.38%,73~168 h 时效下暴雨预报 TS 评分则提升了 43.53%(图 7b)。对大暴雨的预报,风清模型 TS 评分较 ECMWF 模式偏低,说明仍需提升其对于极端过程的预报能力。此外,大雨预报的 Bias 评分,风清模型较 ECMWF 模式偏高,基本均超过 1,在长期时效下 Bias 评分甚至高达 5,说明风清模型对大雨的空报偏多。

#### 3.2.2 暴雨个例预报检验

2024 年全国共出现 41 次大范围降水过程,根据行业标准《降雨过程强度等级》(QX/T 341—2016;中国气象局,2016)可知,其中 18 次达到强等级,达到特强等级的暴雨过程共有 4 次。前文提及的台风“格美”,便是引发其中一次特强等级暴雨过程的“主角”。7 月下旬,台风“格美”本体为浙江、江

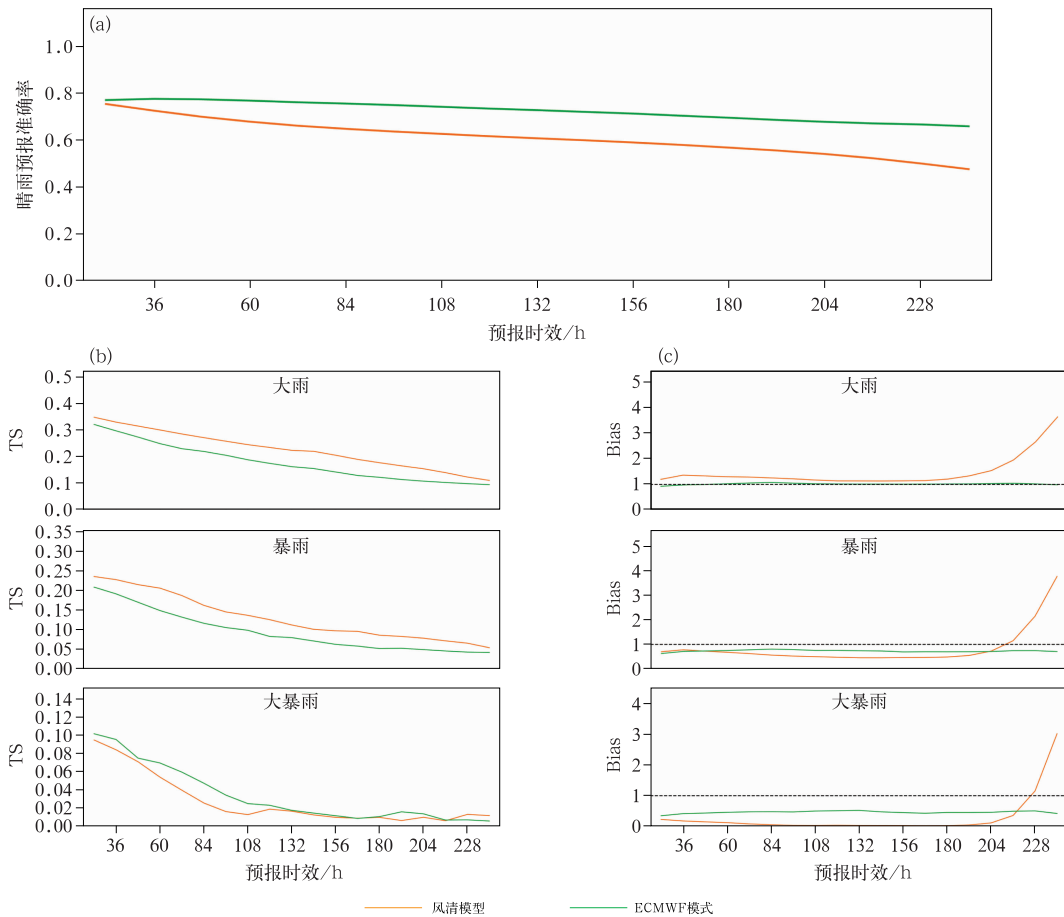


图 7 风清模型和 ECMWF 模式对 2024 年 24 h 降水的(a)晴雨预报准确率, (b,c)大雨、暴雨及大暴雨预报(b)TS 评分和(c)Bias 评分

Fig. 7 Verification of 24 h precipitation forecasts by Fengqing and ECMWF models in 2024 (a) accuracy rate of rain/no-rain, (b) TS score and (c) Bias score for heavy rain, rainstorm and severe rainstorm

西、湖南等地带来了强降雨,其外围云系和远距离输送效应还造成了山东、辽宁、吉林等地的暴雨。

图 8 展示了此次过程中区域平均降水最强日(7月 28 日 08 时)的 24 h 降水量预报检验结果。从观测降水空间分布可知,受台风影响,当日降水呈弧状分布,华南—江南地区(广东、湖南)、华东地区(安徽、山东)以及东北地区(辽宁)为 3 个强降水中心。值得注意的是,风清模型和 ECMWF 模式在 36 h 时效下预报雨带基本特征方面都有不错的表现,风清模型预报的降水空间分布与观测的空间相关性达到 0.63,ECMWF 模式为 0.49。在对强降水中心的预报上,二者都表现良好,风清模型优势尤为突出;观测的 24 h 降水量超过 50 mm 的站点有 179 个,而风清模型和 ECMWF 模式预报分别命中 74 个和 60 个;但从极端性的角度来说,观测的最大降水量为 313.1 mm,ECMWF 模式预报可达 236.5 mm,而风清模型预报的最大降水量仅为 168.9 mm,远低于观测结果。对于此次过程的暴雨和大暴雨预报评分,风清模型分别可达 0.27 和 0.18,显著高于 ECMWF 模式。通过定量检验可知,0~72 h 预报的该日暴雨、大暴雨 TS 评分均较 ECMWF 模式分别偏高 18.0%、55.4%(表略),说明风清模型对此次降水过程预报效果较好,其在降水预报上展现出的优势,可能与前文所述其在台风移动路径预报方面的出色表现相关。

除台风降水外,梅雨锋降水也是预报关注重点之一。2024 年 6 月 21 日至 7 月 2 日,受高空槽、低涡切变和低空急流共同影响,我国长江中下游、广西和贵州等地出现持续强降水过程,此次过程也是年度 4 次特强等级暴雨过程之一。

图 9 给出此次过程 6 月 27 日 08 时起报的 36 h 时效的 24 h 降水量其对应的环流预报结果。从图中可以看出,风清模型和 ECMWF 模式均可以较好预报东西向雨带的基本特征。值得注意的是,观测中雨带东部的暴雨-大暴雨区主要位于湖北东部到安徽南部一带,风清模型对此预报能力较好,强降水落区位置基本与观测相当,但 ECMWF 模式预报的强降水中心则明显偏北。通过定量检验可知,风清模型和 ECMWF 模式对暴雨(大暴雨)的预报 TS 评分分别为 0.559(0.286)和 0.517(0.107),即 36 h 时效下风清模型均更优,且在其他时效下的预报 TS 评分也存在优势(表略)。上述结果展示了风清模型对强降水较好的预报能力,这与其针对性考虑降水分布样本不均的问题,增加了对强降水预报的训练权重有关。

从环流预报的角度来看,风清模型和 ECMWF 模式均可预报此次过程中高空槽、低空急流的特征。对比预报结果与实际观测的偏差可知,ECMWF 模式对雨带预报偏北的主要原因在于其对河南南部上空的低涡强度预报显著偏强,风速预报亦偏大。

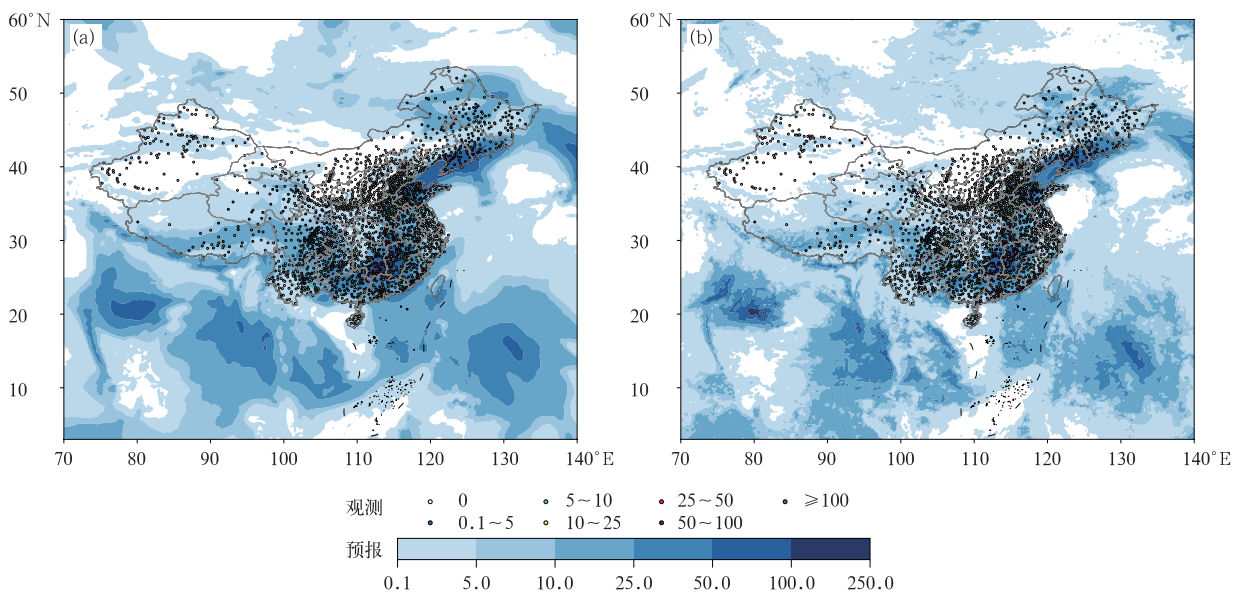
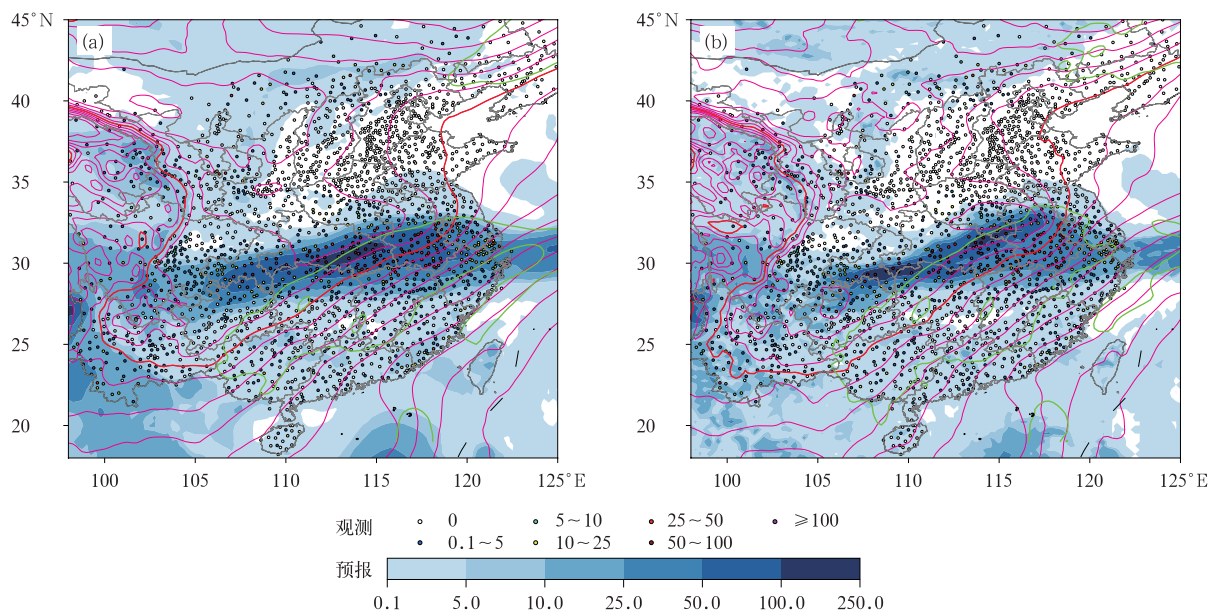


图 8 (a)风清模型,(b)ECMWF 模式 2024 年 7 月 26 日 20 时起报的 36 h 预报时效的 24 h 降水量(填色)与观测降水(圆点)的比较(单位:mm)

Fig. 8 Comparison between the observed precipitation (dot) and the 24 h precipitation forecasts (colored) by (a) Fengqing and (b) ECMWF models with 36 h lead time initialized at 20:00 BT 26 July 2024 (unit: mm)

为进一步了解风清模型与 ECMWF 模式对雨带特征的预报能力,针对 2024 年强和特强等级的暴

雨过程(共 22 次)进行基于对象的诊断评估(简称 MODE 空间检验),图 10 给出其中可识别的 19 次



注:粉色等值线为 850 hPa 位势高度,红色等值线为 1440 gpm 线,绿色等值线为  $12 \text{ m} \cdot \text{s}^{-1}$  风速。

图 9 (a)风清模型,(b)ECMWF 模式 2024 年 6 月 27 日 08 时起报的 36 h 预报时效的 24 h 降水量(填色,单位:mm)及相应的环流场(彩色等值线)预报与观测降水(圆点,单位:mm)的比较

Fig. 9 Comparison between the observed precipitation (dot, unit: mm) and the 24 h precipitation forecasts (colored, unit: mm) and its circulation field (colored contour) forecasts by (a) Fengqing and (b) ECMWF models with 36 h lead time initialized at 08:00 BT 27 June 2024

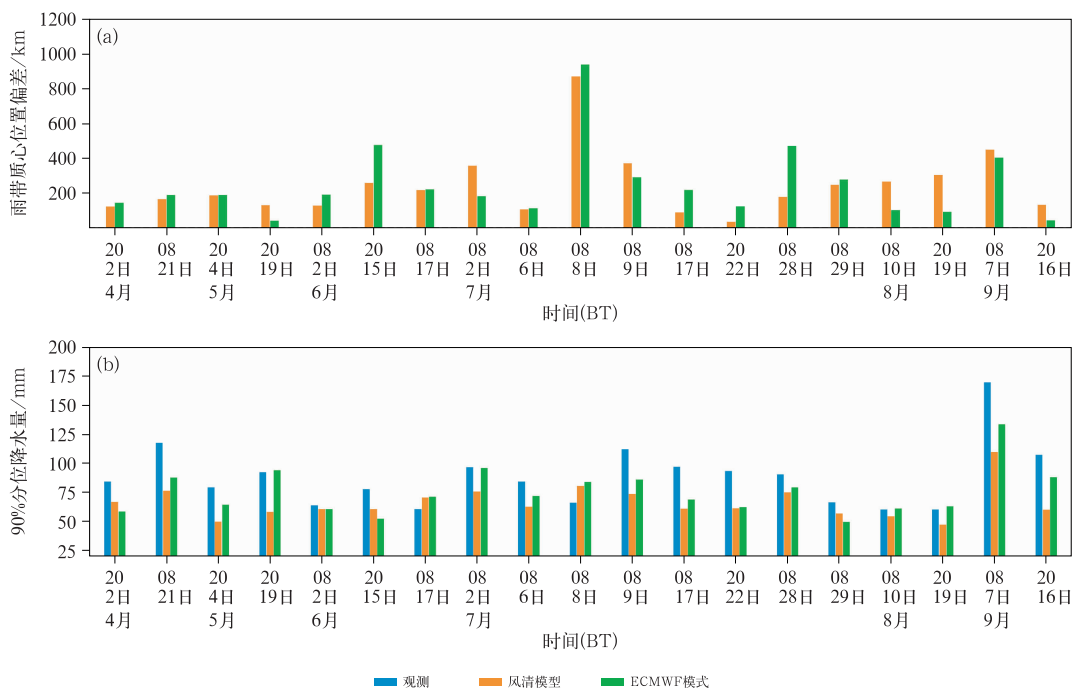


图 10 风清模型、ECMWF 模式对 2024 年 19 次强和特强等级暴雨过程 24 h 降水雨带(a)质心位置偏差和(b)90%分位降水量值分布的 MODE 检验结果与观测的对比

Fig. 10 Comparison of MODE test results by Fengqing and ECMWF models with observations for (a) the centroid position deviation and (b) 90th percentile precipitation values of the 24 h precipitation belts during 19 severe and extremely severe rainstorm events in 2024

暴雨过程的24 h降水雨带特征检验结果。对于大雨以上的雨区来说,风清模型48 h时效内预报的雨带质心位置预报偏差在38.5~870.1 km,ECMWF模式的48 h时效内预报雨带质心位置预报偏差在45.1~939.4 km;二者预报的雨带质心位置平均误差分别为245.3 km和250.8 km,且19次过程中风清模型预报的雨带质心位置偏差小于ECMWF模式预报结果的有12次。通过对雨带内90%分位24 h降水量(用于表征极端强降水)进行检验可知,风清模型和ECMWF模式预报的结果较观测分别偏低23.15 mm和13.08 mm。

总体来说,从个例和统计检验结果均可发现风清模型对雨带位置的预报较好,但对极端降水强度的预报偏弱。

## 4 结论和讨论

本文基于观测,从预报准确率、偏差分布及典型天气过程等方面,对2024年风清模型在中国及周边区域的预报性能开展系统性检验,并与ECMWF全球高分辨率模式预报结果进行了对比评估。主要结论如下:

(1)风清模型对850 hPa温度、风速及地面要素的预报效果明显优于ECMWF模式结果,尤其是在青藏高原及其周边的复杂地形区域优势更为突出,其中2 m气温预报的RMSE较ECMWF模式的最大改进幅度达37.66%。此外,500 hPa位势高度场的有效预报时效可达10 d。该模型在抑制中长期时效下误差快速增长、维持大尺度环流演变特征方面的优势可能与其采用的物理约束方案有关。

(2)在典型天气过程的预报方面,由于风清模型在训练阶段针对性地引入了与极端过程相关的物理约束与权重优化策略,因此对台风移动路径、暴雨落区的预报能力较强,相应的预报误差在大部分时效下较ECMWF模式均偏小。但风清模型仍存在对灾害性天气过程极端性预报不足的问题,具体表现为大暴雨整体TS评分较低、台风强度预报偏弱。

(3)随着预报时效的增加,风清模型预报活跃度逐渐下降,说明其对中长期时效下天气变化,尤其是中小尺度系统变化的预报能力不足。这可能与模型训练方法和不同尺度系统固有的可预报性有关。

风清模型是中国气象局发布的首个人工智能全球中短期预报系统。该模型通过引入多尺度隐空间

投影及能量守恒约束等设计,实现了数据驱动与动力学约束的有机结合。检验结果表明,在保持物理一致性的基础上,该模型显著改善了我国复杂地形区及中长期时效下关键要素的预报质量,这为业务预报提供了新的参考。同时检验结果也指出,风清模型在灾害性天气极端性预报方面仍存在局限。其预报的台风中心强度、暴雨强度存在系统性偏弱的问题。因此,在实际业务应用中,风清模型的预报产品可以作为环流形势、降水落区判断的重要参考,但涉及极端性的定量预报时,预报员仍需结合传统数值预报结果进行综合判断。需要指出的是,本文仅对2024年风清模型预报结果进行了检验评估,未来还需利用更长时间的预报数据进行检验,以揭示模型在不同气候背景下的偏差特征。此外,未来可以引入更多能反映极端性的指标(如极端高/低温日数、极端降水强度),考察人工智能模型对极端高温、寒潮和暴雨的预报效果,从而更全面地分析模型在实际业务预报中的适用性。

## 参考文献

- 陈鹤,蔡荣辉,陈静静,等,2022. 基于深度学习方法的温度预报技术应用与评估[J]. 气象,48(11):1373-1383. Chen H, Cai R H, Chen J J, et al, 2022. Application and evaluation of temperature forecast based on deep learning method[J]. Meteor Mon, 48(11):1373-1383(in Chinese).
- 黄小猛,林岩奎,熊巍,等,2024. 数值预报AI气象大模型国际发展动态研究[J]. 大气科学学报,47(1):46-54. Huang X M, Lin Y L, Xiong W, et al, 2024. Research on international developments of AI large meteorological models in numerical forecasting[J]. Trans Atmos Sci, 47(1):46-54(in Chinese).
- 刘会军,吴启树,危国飞,等,2024. 数值模式降水预报OTS订正法的实现技术[J]. 大气科学,48(5):1891-1900. Liu H J, Wu Q S, Wei G F, et al, 2024. Implement technology of optimal threat score correction method for numerical model precipitation forecast[J]. Chin J Atmos Sci, 48(5):1891-1900(in Chinese).
- 马雷鸣,2020. 天气预报中的人工智能技术进展[J]. 地球科学进展,35(6):551-560. Ma L M, 2020. Development of artificial intelligence technology in weather forecast[J]. Adv Earth Sci, 35(6):551-560(in Chinese).
- 梅婵娟,王秀明,李蕊,等,2025. 渤海湾飑线入海增强机制的个例研究[J]. 气象,51(9):1061-1071. Mei C J, Wang X M, Li R, et al, 2025. A case study on the enhancement mechanism of a squall line over the Bohai Bay after moving into the sea[J]. Meteor Mon, 51(9):1061-1071(in Chinese).
- 邱孙俊杰,张芳华,2025. 2025年6月大气环流和天气分析——入梅以来南方多地出现极端强降水[J]. 气象,51(9):1146-1156. Qiu S J J, Zhang F H, 2025. Atmospheric circulation and weather

- analysis in June 2025; persistent extreme heavy rainfall in southern China since Meiyu onset[J]. *Meteor Mon*, 51(9):1146-1156 (in Chinese).
- 沈学顺,陈起英,孙健,等,2021.中央气象台全球中期数值预报业务系统的发展[J].*气象*,47(6):645-654. Shen X S, Chen Q Y, Sun J, et al, 2021. Development of operational global medium-range forecast system in National Meteorological Centre[J]. *Meteor Mon*, 47(6):645-654 (in Chinese).
- 盛杰,金荣花,张小雯,等,2025.临近气象预报大模型“风雷”V1 版本检验及个例评估[J].*气象*,51(4):389-399. Sheng J, Jin R H, Zhang X W, et al, 2025. Verification and case evaluation of the “Fenglei” V1 meteorological nowcasting model [J]. *Meteor Mon*, 51(4):389-399 (in Chinese).
- 孙一听,张楠,2025.天津机场一次极端雷暴大风过程的多尺度特征分析[J].*气象*,51(9):1101-1112. Sun Y X, Zhang N, 2025. Multi-scale characteristics analysis of an extreme thunderstorm gale event at Tianjin Airport[J]. *Meteor Mon*, 51(9):1101-1112 (in Chinese).
- 杨绚,代刊,朱跃进,2022.深度学习技术在智能网格天气预报中的应用进展与挑战[J].*气象学报*,80(5):649-667. Yang X, Dai K, Zhu Y J, 2022. Progress and challenges of deep learning techniques in intelligent grid weather forecasting[J]. *Acta Meteor Sin*, 80(5):649-667 (in Chinese).
- 中国气象局,2016.降雨过程强度等级:QX/T 341—2016[S].北京:气象出版社. China Meteorological Administration, 2016. The intensity grade of rainfall process; QX/T 341—2016[S]. Beijing: China Meteorological Press (in Chinese).
- 周康辉,代刊,曾沁,等,2025.全球人工智能治理现状及气象领域人工智能治理的挑战[J].*气象学报*,83(1):168-177. Zhou K H, Dai K, Zeng Q, et al, 2025. Current status of artificial intelligence governance and its challenges in the meteorological community[J]. *Acta Meteor Sin*, 83(1):168-177 (in Chinese).
- Bi K F, Xie L X, Zhang H H, et al, 2023. Accurate medium-range global weather forecasting with 3D neural networks[J]. *Nature*, 619(7970):533-538.
- Bojesomo A, Al-Marzouqi H, Liatsis P, 2023. A novel transformer network with shifted window cross-attention for spatiotemporal weather forecasting[J]. *IEEE J Sel Top Appl Earth Observ Remote Sens*, 17:45-55.
- Bouallègue Z B, Clare M C A, Magnusson L, et al, 2024. The rise of data-driven weather forecasting: a first statistical assessment of machine learning-based weather forecasts in an operational-like context[J]. *Bull Am Meteor Soc*, 105(6):E864-E883.
- Chen K, Han T, Ling F H, et al, 2025. The operational medium-range deterministic weather forecasting can be extended beyond a 10-day lead time[J]. *Commun Earth Environ*, 6(1):518.
- Chen L Y, Han B C, Wang X S, et al, 2023. Machine learning methods in weather and climate applications: a survey[J]. *Appl Sci*, 13(21):12019.
- Ebert-Uphoff I, Hilburn K, 2023. The outlook for AI weather prediction[J]. *Nature*, 619(7970):473-474.
- Hersbach H, Bell B, Berrisford P, et al, 2020. The ERA5 global reanalysis[J]. *Quart J Roy Meteor Soc*, 146(730):1999-2049.
- Keisler R, 2022. Forecasting global weather with graph neural networks[PP]. arXiv. <https://doi.org/10.48550/arXiv.2202.07575>.
- Lam R, Sanchez-Gonzalez A, Willson M, et al, 2023. Learning skillful medium-range global weather forecasting[J]. *Science*, 382(6677):1416-1421.
- Liu C C, Hsu K, Peng M S, et al, 2024. Evaluation of five global AI models for predicting weather in Eastern Asia and Western Pacific[J]. *npj Climate Atmos Sci*, 7(1):221.
- Nguyen T, Brandstetter J, Kapoor A, et al, 2023. ClimaX: a foundation model for weather and climate[C]//Proceedings of the 40th International Conference on Machine Learning. Honolulu: PMLR.
- Niu Z Y, 2025. ML-physical fusion models are accelerating the paradigm shift in operational typhoon forecasting[PP]. arXiv. <https://arxiv.org/abs/2503.00424>.
- Olivetti L, Messori G, 2024. Do data-driven models beat numerical models in forecasting weather extremes? a comparison of IFS HRES, Pangu-Weather, and GraphCast[J]. *Geosci Model Dev*, 17(21):7915-7962.
- Pathak J, Subramanian S, Harrington P, et al, 2022. FourCastNet: a global data-driven high-resolution weather model using adaptive fourier neural operators [PP]. arXiv. <https://doi.org/10.48550/arXiv.2202.11214>.
- Radford J T, Ebert-Uphoff I, Stewart J Q, et al, 2025. Accelerating community-wide evaluation of AI models for global weather prediction by facilitating access to model output[J]. *Bull Amer Meteor Soc*, 106(1):E68-E76.
- Rasp S, Dueben P D, Scher S, et al, 2020. WeatherBench: a benchmark data set for data-driven weather forecasting [J]. *J Adv Model Earth Syst*, 12(11):e2020MS002203.
- Rasp S, Thuerey N, 2021. Data-driven medium-range weather prediction with a resnet pretrained on climate simulations: a new model for WeatherBench[J]. *J Adv Model Earth Syst*, 13(2):e2020MS002405.
- Schultz M G, Betancourt C, Gong B, et al, 2021. Can deep learning beat numerical weather prediction? [J]. *Phil Trans R Soc A*, 379(2194):20200097.
- Wu H X, Hu T G, Luo H K, et al, 2023. Solving High-dimensional PDEs with latent spectral models[PP]. arXiv. <https://doi.org/10.48550/arXiv.2301.12664>.
- Xu D S, Lu Z B, Leung J C H, et al, 2025. AI models still lag behind traditional numerical models in predicting sudden-turning typhoons[J]. *Sci Bull*, 2025, 70(17):2705-2708.