安徽农业大学学报, 2026, 53(4): 673-688
Journal of Anhui Agricultural University [DOI] 10.13610/j.cnki.1672-352x.20260918.002

吴德平，王佳颖，陈冠霖，等. GFS、CMA-GFS、WRF 模式与 AI 气象大模型对台风“麦德姆”预报能力对比分析[J]. 安徽农业大学学报, 2026, 53(4): 673-688.

Wu Deping, Wang Jiaying, Chen Guanlin, et al. Comparative analysis of GFS, CMA-GFS, WRF and AI weather models for Typhoon Matmo forecast[J]. Journal of Anhui Agricultural University, 2026, 53(4): 673-688.

GFS、CMA-GFS、WRF 模式与 AI 气象大模型对台风“麦德姆”预报能力对比分析

吴德平[1,2]，王佳颖[1]，陈冠霖[1]，唐　源[1]，谭淇昌[3]，张　宇[1*]

(1. 广东海洋大学海洋与气象学院，粤西近海海洋气象灾害理论与预测预报应用研究重点实验室，广东 湛江 524088；
2. 广东省湛江市气象局，广东 湛江 524001；3. 成都信息工程大学大气科学学院，四川 成都 610225)

摘　要：【目的】评估 2025 年台风“麦德姆”（2521）路径、强度及结构预报中多模式性能差异，为业务预警与模式改进提供参考。【方法】基于中国气象局（CMA）最佳路径数据与欧洲中期天气预报中心再分析数据（ERA5），对比两类全球数值模式、两类区域降尺度模式及两种人工智能（AI）气象大模型的预报结果，采用平均绝对误差（MAE）和均方根误差（RMSE）定量评估 24、48、60 与 96 h 预报时效的路径与强度误差，并结合暖心、动力结构及大尺度环流诊断分析误差机制。【结果】在 24～60 h 路径预报中，FuXi 模型精度最优，48 h MAE 为 25 km；96 h RMSE 增至 417.4 km。CMA-GFS（China Meteorological Administration Global Forecast System ）模型短时效偏差较大，但随预报时效延长误差增长较缓；GFS 模型（Global Forecast System）短期预报较好，长时效易偏移。WRF-GFS（Weather Research and Forecasting model driven by the Global Forecast System）与 WRF-CMA-GFS（Weather Research and Forecasting model driven by the China Meteorological Administration Global Forecast System）短时效误差偏大，但长时效稳定性突出。在强度预报方面，Pangu 和 FuXi 模型全程显著低估强度，无法刻画近海快速降压过程；WRF-CMA-GFS 对快速增强与峰值强度的模拟最优，且预报质量显著依赖初始场。【结论】GFS 与 CMA-GFS 在路径与强度预报中表现出显著差异：路径预报上二者呈短–长时效能力反转特征，即 GFS 短期占优但长时效误差骤增，CMA-GFS 短期偏差明显但长时效最稳定；强度预报表现则与该特征相反。WRF 模式长时效预报性能突出，且质量显著依赖初始场。本文通过“麦德姆”这一近海增强个例的多模式对比，揭示了不同模式与 AI 模型在不同时效中的差异化表现及结构模拟的关键作用，研究结果可为台风多模式集成优选、时效策略制定提供量化依据，为 AI 模型融入物理约束、区域模式优化初始场提供改进方向，对提升近海增强台风预报业务能力具有一定应用价值。

关键词：台风“麦德姆”；数值预报模式；AI 气象大模型；路径预报；强度预报

中图分类号：P457.8　　**文献标志码：**A　　**文章编号：**1672−352X (2026)04−0673−16

Comparative analysis of GFS, CMA-GFS, WRF and AI weather models for Typhoon Matmo forecast

Wu Deping[1,2], Wang Jiaying[1], Chen Guanlin[1], Tang Yuan[1], Tan Qichang[3], Zhang Yu[1]

(1. Western Guangdong Key Laboratory of Marine Meteorological Disaster Theory and Application, College of Ocean and Meteorology, Guangdong Ocean University, Zhanjiang 524088, China; 2. Zhanjiang Meteorological Bureau of Guangdong Province, Zhanjiang 524001, China; 3. School of Atmospheric Sciences, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: [Objective] This study evaluates the performance differences among multiple models in forecasting the track, intensity, and structure of Typhoon Matmo (2521), aiming to provide references for operational warning and model improvement. **[Method]** Based on the China Meteorological Administration (CMA) best track data and European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis data (ERA5), this study compared the

收稿日期: 2026-03-24；**录用日期：**2026-05-27；**网络出版日期：**2026-09-20

基金项目：国家自然科学基金项目（42375159）；粤西海洋气象灾害理论与应用研究重点实验室开放项目(2025KSYS009)；广东海洋大学大学生创新训练项目（CXXL2026012）

共同第一作者：吴德平，高级工程师，E-mail：zjwdp@126.com；王佳颖，本科生，E-mail：wjy191022_2024@qq.com

* **通信作者：**张　宇，博士，教授，博士生导师，E-mail：zhangyu2@gdou.edu.cn

forecast results of two global numerical models, two regional downscaling models, and two artificial intelligence (AI) meteorological large models. The mean absolute error (MAE) and root mean square error (RMSE) were used to quantitatively evaluate track and intensity errors at 24-hour, 48-hour, 60-hour, and 96-hour forecast horizons. Error mechanisms were also diagnosed through analyses of thermal structure, dynamic structure, and large-scale circulation. **[Result]** For 24-60-hour track forecasts, the FuXi model demonstrated the highest accuracy, with a 48-h MAE of 25 km; however, the 96-hour RMSE increased to 417.4 km. The China Meteorological Administration Global Forecast System (CMA-GFS) model exhibited larger errors in the short-range forecast period, but its error growth rate slowed as the forecast horizon extended; conversely, the Global Forecast System (GFS) performed well in the short term but tended to deviate in the long term. The Weather Research and Forecasting (WRF) models driven by the Global Forecast System (WRF-GFS) and Weather Research and Forecasting model driven by the China Meteorological Administration Global Forecast System (WRF-CMA-GFS) models had relatively large errors in the short range forecast period but demonstrated outstanding stability in the long rangeforecast period. Regarding intensity forecasts, Pangu and FuXi significantly underestimated intensity throughout the entire track and failed to capture rapid pressure drops near the coast. WRF-CMA-GFS performed best in simulating rapid intensification and peak intensity, although its forecast quality was significantly dependent on initial conditions. **[Conclusion]** GFS and CMA-GFS exhibit significant differences in track and intensity forecasts. Regarding tracks, the two models show a reversal of short-term versus long-term capabilities: GFS performs better in the short term but shows a sharp increase in errors over the long term, while CMA-GFS exhibits significant short-term deviations but is the most stable over the long term. Intensity forecasts, however, show the opposite pattern. The Weather Research and Forecasting WRF model demonstrates outstanding long-range forecasting performance, with intensity forecasts generally outperforming track forecasts, though forecast quality remains significantly dependent on initial conditions. Through a multi-model comparison of Typhoon Matmo—a rapid offshore intensification case-this study reveals the differentiated performance of various models across different time scales and highlights the critical role of structural simulation. These findings provide quantitative evidence for optimizing typhoon multi-model ensembles and formulating time-scale-dependent strategies, offer directions for integrating physical constraints into AI models and optimizing initial conditions for regional models, and hold practical value for enhancing operational forecasting capabilities for rapidly intensifying typhoons near the coast.

Keywords: Typhoon Matmo; numerical weather prediction models; AI weather models; track forecasting; intensity forecasting

热带气旋是地球上最具破坏性的自然灾害系统之一，其伴随的狂风、暴雨和风暴潮常对沿海地区的生命财产安全与社会经济稳定构成严重威胁[1-2]。广东省地处我国大陆最南端，南临南海且海岸线绵长（约 4 114 km），粤西雷州半岛等沿海区域地势低平、无地形屏障，直面南海北部开阔海域，是西北太平洋台风西进登陆的核心通道。加之南海北部常年 26 ℃以上的海表温度为台风提供充足热力支撑，使得广东成为我国受台风影响最频繁、强度最显著的省份之一。近年来，受气候变化影响，台风的活动规律与强度特征更趋复杂，因此，实现对其路径与强度的高精度预报已成为提升防灾减灾能力的迫切需求[3]。

当前，台风预报主要依赖传统数值天气预报模式与人工智能（artificial intelligence，AI）气象大模型。传统数值天气预报（numerical weather prediction，NWP）模式历经数十年迭代，已形成“全球-区域高分辨率”业务预报体系，在台风路径与强度预报方面持续取得进展。在全球业务模式方面，NCEP-GFS 作为国际主流全球确定性预报系统，采用半隐式半拉格朗日动力框架与多物理过程参数化方案，并通过卫星与地面观测资料同化持续优化；该系统在西北太平洋台风路径预报中性能稳定，但对近海快速增强台风的强度预报仍存在系统性偏差，且对副热带高压与引导气流模拟的不确定性易导致长时效路径误差增大[4]。我国自主研发的 CMA-GFS 全球模式实现了关键技术跨越，先后完成四维变分同化（4DVar）、台风涡旋初始化（BDA）、集合-变分同化（En4DVar）等核心升级；其水平分辨率达 0.125°，卫星资料同化占比超 75%，在东亚与南海台风预报中路径误差显著降低，长时效稳定性突出，对复杂环流背景下台风路径调控能力显著提升[5-7]。在区域高分辨率模式方面，WRF 模式凭借灵活的物理方案配置与嵌套网格技术，成为台风精细结构模拟的主流工具；该模式以全球模式为驱动场开展动力降尺度，可有效刻画台风眼墙、暖心

结构与次级环流，显著改善强度与降水预报，在近海快速增强台风模拟中优势明显。

以深度学习为核心的 AI 气象大模型重构了全球气象预报范式，Pangu-Weather 与 FuXi 是其中的代表。Pangu-Weather 基于三维地球专用 Transformer（3DEST）架构，突破了传统神经网络在气象场时空建模方面的局限，以 0.25°分辨率实现了中期全球气象预报精度超越业务数值天气预报（NWP），预报速度提升万倍以上；该模型在大尺度环流与台风路径预报中性能比肩 ECMWF-IFS，但对台风内核结构刻画不足，强度预报存在显著低估[8]。FuXi 采用级联式 U-Transformer 架构，分时段优化短期、中期及长期预报模块，支持集合预报与不确定性量化，在东亚区域与台风预报中表现优异，路径预报精度稳步提升；此外，FuXi 进一步实现了端到端机器学习数据同化与循环预报，在观测稀疏区域性能超越 ECMWF-HRES，为台风业务预报提供了新支撑[9-10]。尽管 AI 模型在大尺度环流与路径预报中潜力巨大，但受数据驱动机制与“黑箱”特性限制，普遍存在台风强度系统性低估、物理场过度平滑以及暖心与眼墙等中尺度结构刻画不足等问题，其物理一致性与极端事件预报可靠性仍需提升[11-12]。

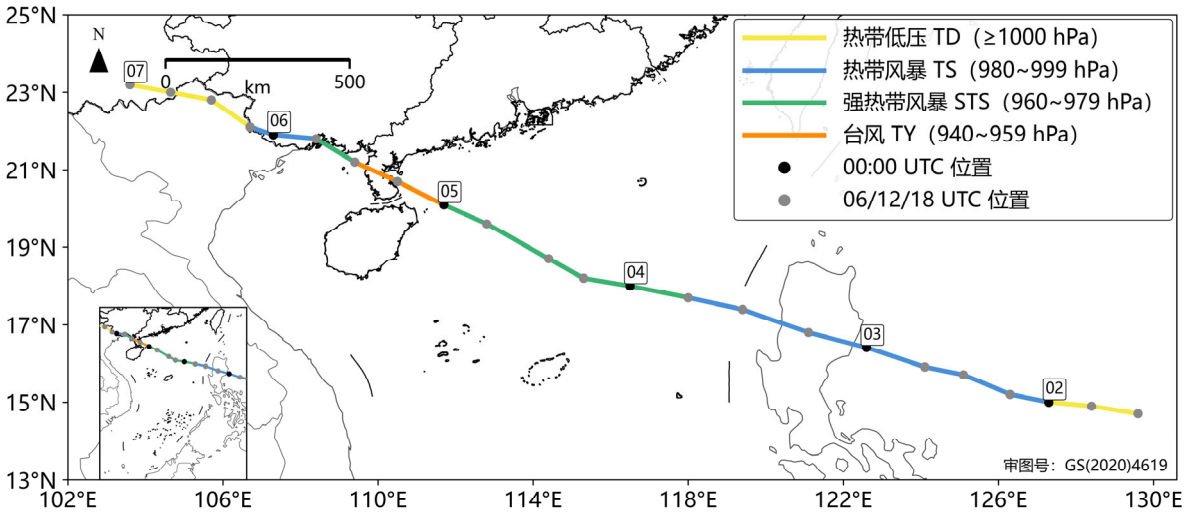


图 1　台风“麦德姆”实时路径及强度

Figure 1　Real-time track and intensity of Typhoon Matmo

在此背景下，以 2025 年台风“麦德姆”为典型个例开展研究。如图 1 所示，该台风于 10 月初在菲律宾以东洋面生成，向西偏北移动并穿越吕宋岛，进入南海后迅速增强为强台风，10 月 5 日下午以强台风级在广东徐闻沿海登陆，成为 2025 年登陆中国的强台风之一，也是 1951 年以来同期登陆广西的最强台风。其具有路径稳定、近海加强、风雨强度大等特点，具备开展多模式对比检验的典型科学价值。其路径稳定、移动趋势清晰，有利于客观识别不同模式的路径偏差特征；同时，该台风在登陆前经历增强，对数值模式 AI 模型均构成较强预报挑战，能够有效体现两类模型在强度预报、结构模拟及大尺度环流响应方面的优势与不足。对菲律宾、中国、越南等多国造成严重风雨影响和灾害损失，截至 2025 年 11 月 11 日，已在我国粤桂琼滇等地导致直接经济损失 219.28 亿元[13]。因此，本文选取“麦德姆”作为研究对象，通过集成多源数值预报产品，系统开展其路径与强度的多模式统计检验与误差诊断，旨在揭示全球数值模式、区域降尺度模式 AI 气象大模型在南海近海增强台风的路径、强度及结构预报差异与误差机理，明确各模型在短、中、长时效的性能特征，为台风预报改进与多模式融合提供科学依据。

1　材料与方法

1.1　数据

1.1.1　*观测与再分析数据*　本研究采用中国气象局（China Meteorological Administration，CMA）台风最佳路径数据集作为模型评估的基准参考，该数据集作为台风路径与强度的观测真值，用于评估所有预报模型的路径误差和强度误差。该数据集覆盖西北太平洋及南海区域，核心要素包括台风中心经纬度、2 min 平均近中心最大持续风速及中心最低海平面气压[14]。

本研究采用欧洲中期天气预报中心（ECMWF）发布的第五代全球大气再分析数据（ECMWF Re-

analysis version 5，ERA5），用于分析台风的环境场与精细三维结构。

包括：2025 年 10 月 1 日 00 时至 10 月 6 日 23 时（UTC）逐小时数据，水平范围为 10°～40°N、90°～130°E，垂直层次选取 200、500 和 850 hPa3 个关键等压面，变量涵盖位势高度、温度、相对湿度、水平风场（u、v 分量）及相对涡度。

1.1.2 *模式与模型数据* 为系统对比不同技术路线的预报能力，本研究选取了 6 类具有代表性的预报模型，包括数值预报模式与前沿 AI 气象大模型，并针对 2025 年台风“麦德姆”的特定起报时次（2025 年 10 月 1 日 12 时和 10 月 3 日 12 时）进行对比分析。为确保可比性，将所有模型的预报结果在时间维度上统一为 6 h 间隔，表 1 详细列出了所评估的模型及其相关配置。

表 1 本文所评估的预报模型概览

Table 1 Overview of the prediction models evaluated in this study

模型类型	模型名称	空间分辨率	时间分辨率	初始场/驱动场来源
全球数值模式	GFS	0.25°×0.25°	6 h	NCEP 业务分析场
	CMA-GFS	0.25°×0.25°	3 h	CMA 4DVar 同化系统
区域数值模式	GFS-WRF	9 km×9 km	3 h	GFS 预报场驱动
	CMA-GFS-WRF	9 km×9 km	3 h	CMA-GFS 预报场驱动
AI 气象大模型	Pangu-Weather	0.25°×0.25°	6 h	GFS 初始场
	FuXi	0.25°×0.25°	6 h	GFS 初始场

全球数值模式：GFS（Global Forecast System）由美国国家环境预报中心（National Centers for Environmental Prediction）负责运行与更新，是全球业务预报中应用最广泛的数值模式之一[2]。与之对比，中国气象局自主研发的全球预报系统 CMA-GFS（China Meteorological Administration Global Forecast System）采用先进的四维变分同化（Four-Dimensional Variational Assimilation）与台风初始化方案，显著提升了对东亚及西太平洋地区天气系统的预报能力[6]。

区域高分辨率模式：为探究动力降尺度对台风预报的改进作用，分别以 GFS 和 CMA-GFS 预报场作为驱动场和边界条件，利用中尺度天气预报模式 WRF（Weather Research and Forecasting Model）进行区域动力降尺度模拟，旨在捕捉更精细的台风中尺度结构[15]。WRF 模式版本为 4.2，水平分辨率为 9 km[8]，物理参数化方案选用 Thompson 云微物理方案、Rapid Radiative Transfer Model（RRTM）长波辐射方案、Dudhia 短波辐射方案、Yonsei University（YSU）边界层方案、改进版 MM5 近地面层方案以及 Kain-Fritsch 积云对流参数化方案。

AI 气象大模型：本研究评估了两个前沿的 AI 模型。Pangu 是基于 3D Earth-Specific Transformer（3DEST）架构的全球天气预报模型，通过融入地球几何先验信息并采用分层时间聚合策略减少迭代预报中的误差累积；其空间分辨率 0.25°×0.25°，时间分辨率达 1 h，可实现高效的中期全球天气预报[7]。本文采用 24、6 和 3 h3 个官方预训练版本，未使用 1 h 版本。预报推理遵循长步长优先的贪心分层时间聚合策略，即优先使用 24 h 模型以减少迭代累积误差，不足部分依次使用 6 和 3 h 模型补足。最终预报结果统一输出为 3 h 时间间隔，生成 24、48、60 和 96 h 4 个关键预报时效[7]。FuXi 是基于 U-Transformer 构建的级联式机器学习预报系统，通过串联分别针对 0～5 d（短期）、5～10 d（中期）、10～15 d（长期）优化的 3 个子模型，以 6 h 时间分辨率和 0.25°空间分辨率实现 15 d 全球预报[8]。本研究中 FuXi 输入的初始时次为 10 月 4 日 12 时与 10 月 4 日 18 时，输出的首个时次为 10 月 5 日 00 时。需要说明的是，本文使用的 Pangu 与 FuXi 预报均基于 GFS 初始场，这可能对台风路径预报产生一定影响，但对强度和结构预报影响较小[11,16]。

1.2 方法

为客观衡量各模型的预报性能，本研究以 CMA 最佳路径数据为基准开展台风预报评估。在路径方面，计算各模型 48 和 96 h 预报时效路径的平均绝对误差（MAE）和均方根误差（RMSE）[17]，并分析误差随预报时效的变化规律；在强度方面，以近中心最低海平面气压与最大风速为核心指标，计算各模型预报的平均偏差与均方根误差，综合衡量预报性能。计算公式如下：

平均绝对误差

$$\mathrm{MAE}=\frac{\sum_{i=1}^{N}(d_{oi}-d_{fi})}{N} \tag{1}$$

平均偏差用于衡量预报值相对于观测值的系统性偏高或偏低趋势。其中，d_{oi} 与 d_{fi} 分别为第 i 个样

本的预报值与观测值，N 为总样本数。MAE 对所有误差赋予线性权重，能够直观地反映预报的平均误差大小，其值越小表明预报性能越稳定。

均方根误差

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{i=1}^{N}(d_{oi}-d_{fi})^2}{N}} \tag{2}$$

均方根误差用于衡量预报值与观测值之间的平均偏差幅度。其中，d_{oi} 与 d_{fi} 分别为第 i 个样本的预报值与观测值，N 为总样本数。RMSE 对较大的误差更为敏感，能够有效地反映预报的精确度，其值越小表明预报性能越好。

2　结果与分析

本文系统评估了各预报模型对台风“麦德姆”的路径与强度预报性能。通过对比分析两类人工智能模型（Pangu、FuXi）与 4 类数值模式（全球模式 GFS、CMA-GFS 及其对应的区域模式 WRF-GFS、WRF-CMA-GFS）在 24、48、60 与 96 h 预报时效的表现，揭示了各模型的误差特征及可能成因。

2.1　台风路径预报能力对比分析

对台风“麦德姆”的路径预报评估结果显示，其移动路径整体呈典型的西行北抬特征，各模型在不同预报时效上的表现存在明显差异，不同模式在误差增长速率、路径偏差方向、关键转折点预报及登陆位置模拟方面表现出明显差异。路径误差整体呈现随预报时效增长而增大的趋势（图 2），短时效误差较小、预报较为集中，长时效离散度显著增加，路径偏差明显扩大。

误差整体较小，预报路径与观测较为接近，表现出较好的短期预报一致性（图 3（a））。WRF-GFS、WRF-CMA-GFS 与 CMA-GFS 的路径误差偏大，预报路径偏离观测更为明显，出现系统性偏北特征，在短时效内的预报表现较差。其中 WRF-CMA-GFS 的 RMSE 达 35.1 km，MAE 为 28.6 km，CMA-GFS 的误差也达到 RMSE 42.3 km、MAE 38.8 km，而 FuXi 与 GFS 则能较好地贴近观测路径，RMSE 与 MAE 均控制在 15 km 上下，展现出短时效路径预报的优势。

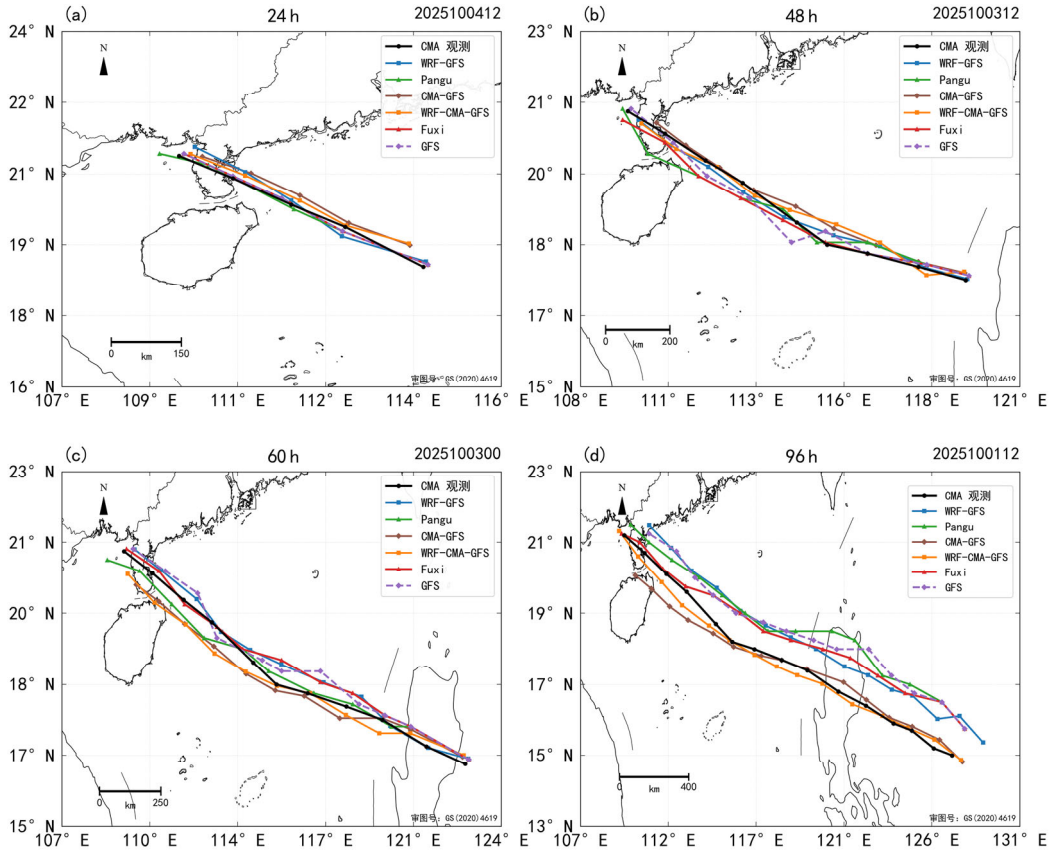


图 2　各数值模式与 AI 模型 24 h(a)、48 h(b)、60 h(c)、96 h(d)预报时效路径预报图

Figure 2　Tropical cyclone track forecasts from different numerical weather prediction models and AI models at forecast lead times of 24 h (a), 48 h (b), 60 h (c), and 96 h (d)

随着预报时效延长至 48 h，各模式与 AI 模型的路径偏差特征逐渐分化，误差水平较 24 h 普遍上升。FuXi 仍保持最优表现，RMSE 为 29.0 km，MAE 为 25.0 km（图 3（b）），路径偏差主要表现为 115°E

以东的轻微偏南趋势（图 2（b））。WRF-GFS 与 GFS 的误差水平相近，RMSE 分别为 31.3 与 32.8 km，路径与观测偏差较小；而 CMA-GFS 与 WRF-CMA-GFS 的误差增幅明显，RMSE 分别升至 50.4 和 44.9 km，路径整体向东南偏移，对台风北抬转折点的预报明显偏晚，导致登陆位置偏南。此时段，Pangu 路径持续偏南，北抬不足的偏差特征已初步显现，误差水平略高于其他数值模式。

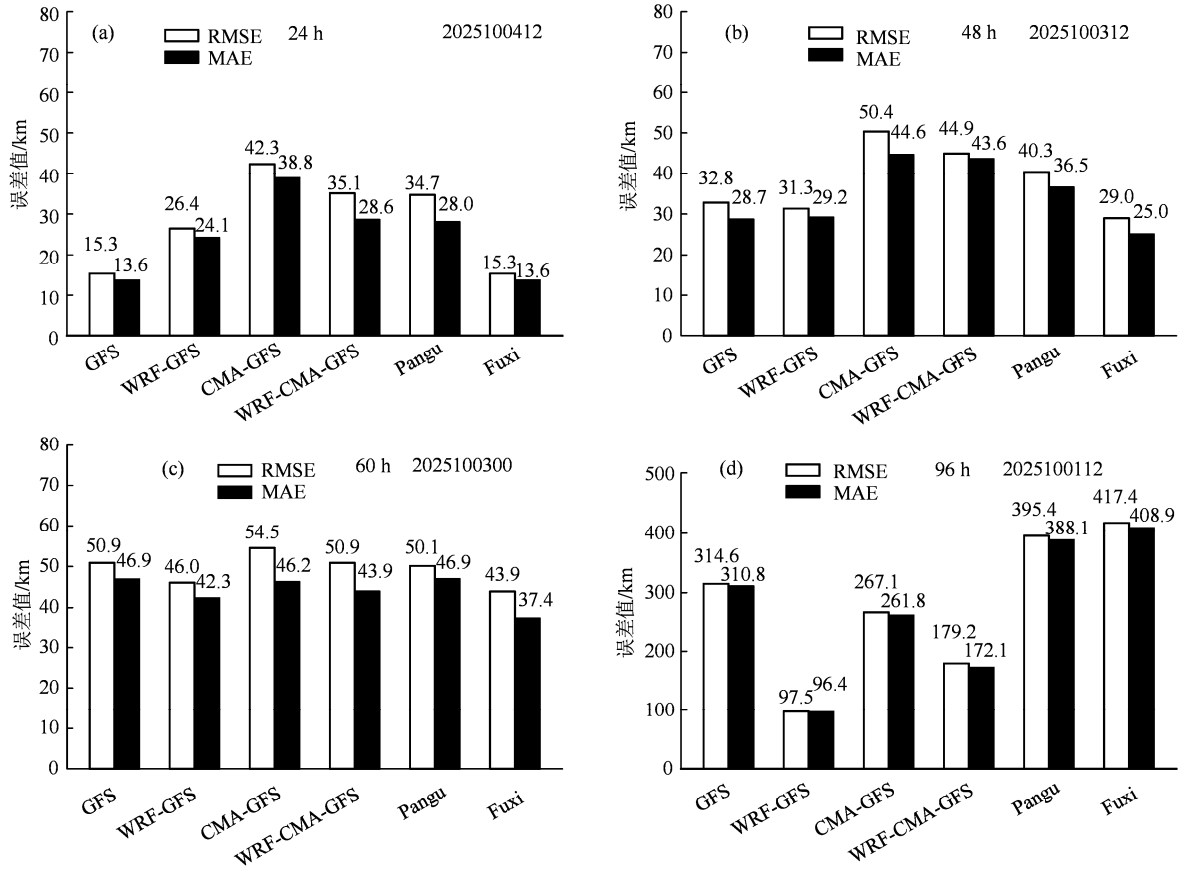


图 3 各数值模式与 AI 模型 24 h(a)、48 h(b)、60 h(c)、96 h(d)预报时效路径误差

Figure 3 Tropical cyclone track forecast errors from different numerical weather prediction models and AI models at forecast lead times of 24 h (a), 48 h (b), 60 h (c), and 96 h (d)

进入 60 h 中期时效，各模式与 AI 模型的预报误差整体差异不大，但误差随预报时效延长的累积效应开始凸显（图 3（c））。FuXi 的误差仍维持在较低水平，RMSE 为 43.9 km，MAE 为 37.4 km，路径在 114°E～119°E 之间显著北偏；CMA-GFS 的误差较大，RMSE 为 54.5 km，MAE 为 46.2 km，对台风北抬位置预报准确，但北抬幅度较小，整体路径偏南且登陆位置偏南（图 2（c））；WRF-CMA-GFS 与 Pangu 的误差水平相近，RMSE 均超过 50 km，路径离散度显著增大。此时段，部分模式已出现对台风移动方向的调整：WRF-CMA-GFS 的路径偏南特征较 48 h 有所减弱，而 GFS 的偏北偏差则进一步扩大，反映出不同模式对大尺度环流演变的预报能力存在差异。

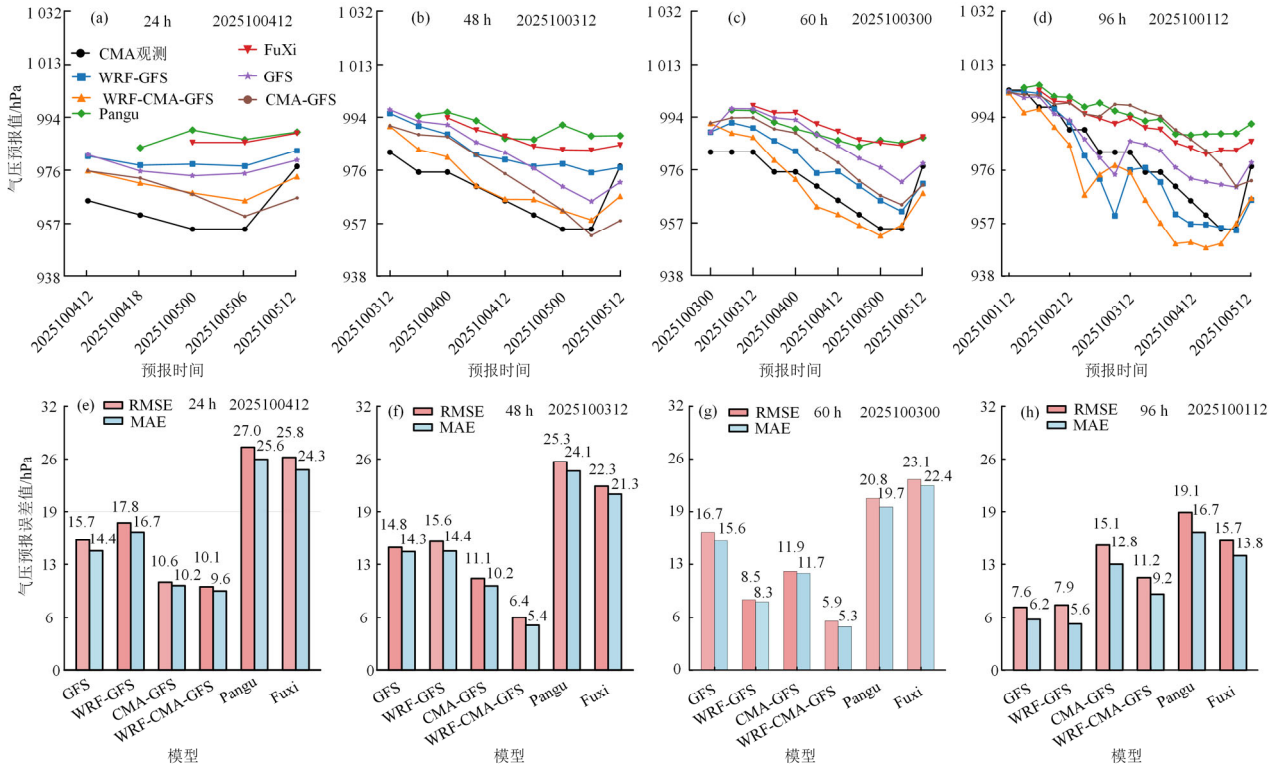
至 96 h 时效，各模式路径预报出现显著分化，误差结构与偏差方向发生明显转变（图 3（d））。除 CMA-GFS 与 WRF-CMA-GFS 路径偏南外，其余模式预报路径均表现为明显偏北，其中 Pangu 与 GFS 的北偏特征最为突出（图 2（d））。WRF-GFS 在长时效内路径误差增幅平缓，预报路径整体接近观测，表现出最优的长时效稳定性；与之相对，GFS 的路径误差急剧增大，呈现显著的系统性偏北特征。AI 模型 FuXi 与 Pangu 的路径误差同样大幅上升，长时效预报稳定性明显下降。在登陆位置预报方面，FuXi 的预报结果与实况最为接近，WRF-GFS 预报登陆位置偏北，CMA-GFS 与 WRF-CMA-GFS 则显著偏南。

综合全时效序列可见，各模式与 AI 模型的路径预报呈现明显的时效依赖性特征。FuXi 在 24～60 h 短至中期时效内路径误差最小，预报稳定性更优，但进入 96 h 长时效后误差增幅显著。WRF-GFS 与 WRF-CMA-GFS 在短时效误差偏大，但随预报时

效延长，误差增长更为平缓，在 96 h 长时效表现最优，稳定性突出。GFS 在短时效表现良好，但长时效误差快速增大，路径呈现系统性偏移。Pangu 在全时效范围内路径误差整体偏高，短时效以偏南为主，长时效转为明显偏北，预报稳定性相对较弱。

2.2　台风强度预报能力对比分析

以台风中心最低海平面气压（MSLP）为评估指标，各模式对台风“麦德姆”强度演变的预报能力随预报时效延长呈现出明显差异，且不同模式在台风快速增强过程、峰值强度及误差演变特征上表现出显著分化（图 4）。在 24～60 h 时效下绝大多数模式与 AI 模型呈现强度预报偏弱特征，96 h 时效下 WRF-GFS 与 WRF-CMA-GFS 预报偏强。



（a）（e）24 h 时效；（b）（f）48 h 时效；（c）（g）60 h 时效；（d）（h）96 h 时效。

图 4　各模式与 AI 模型海平面最低气压预报和误差

Figure 4　Forecast and error distribution of minimum sea level pressure for various numerical models and AI model

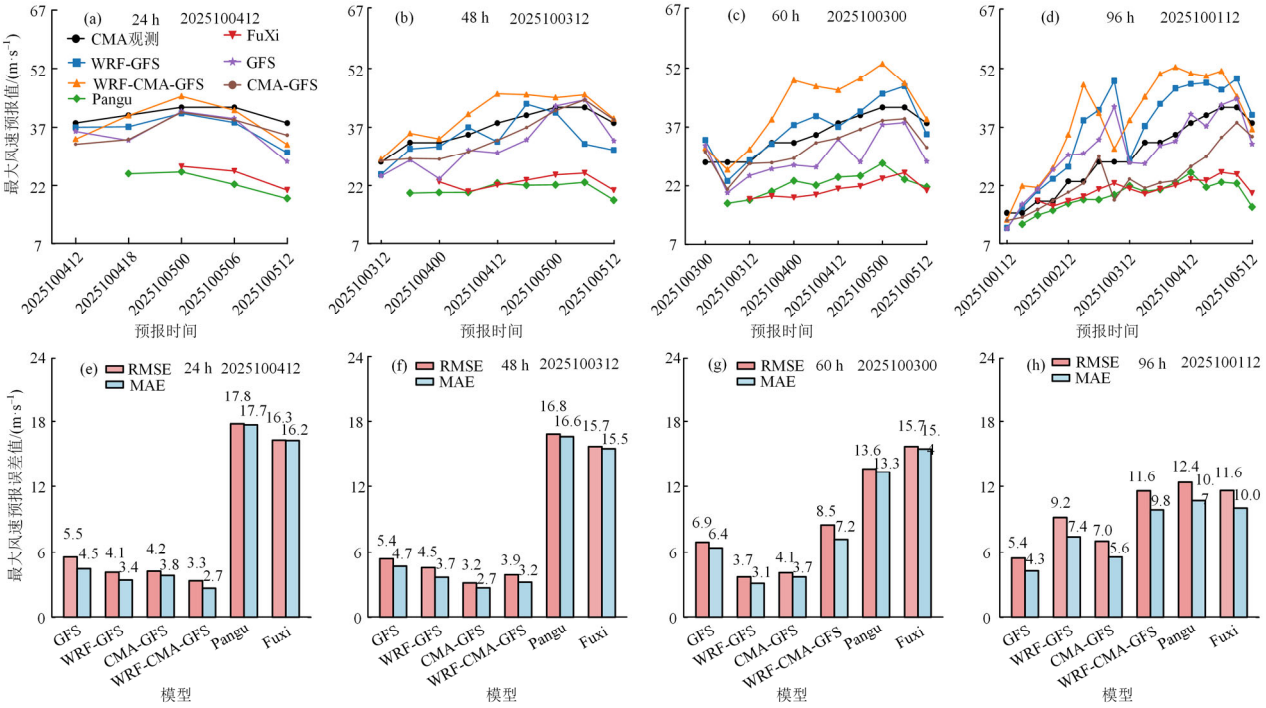
如图 4（a）（e）所示，在 24 h 预报时效内，台风在 10 月 4 日 12 时至 10 月 5 日 06 时已呈现持续增强趋势，中心气压由约 965 hPa 逐步下降，于 10 月 5 日 12 时气压抬升。其中 CMA-GFS 与 WRF-CMA-GFS 对台风增强过程的刻画较为准确，GFS 与 WRF-GFS 也能捕捉到增强过程，但是幅度较小，更加平滑。AI 模型 Pangu 和 FuXi 气压明显偏高，下降过程平缓，无法响应台风初期增强趋势，误差水平显著高于所有数值模式。

在 48 h 台风 MSLP 预报中，区域模式 WRF-CMA-GFS 表现最优（RMSE = 6.4 hPa），预报结果较为可靠；全球模式 CMA-GFS 次之（RMSE = 11.1 hPa）。台风“麦德姆”于 10 月 3 日 12 时至 10 月 5 日 00 时经历增强（图 4（b）），CMA 观测 MSLP 从 980 hPa 降至 950 hPa。数值模式可较好地模拟 MSLP 快速下降趋势，大部分模型的气压最低值出现在 10 月 5 日 06 时，较观测结果略有滞后；而 Pangu 与 FuXi 的 MSLP 曲线近乎平缓，且最低值分别停留在 989 hPa 和 986 hPa，未能有效响应快速增强期的剧烈气压变化。进一步分析表明，AI 模型存在明显的“平滑效应”，其气压场标准差（约 4 hPa）远小于观测值（10.30 hPa）及传统模式（7～14 hPa），倾向于环境背景气压场，无法有效刻画台风内核的极端气压结构。

如图 4（c）（g）所示，在 60 h 预报时效下，各数值模式与 AI 模型间的误差分异最为明显。区域模式 WRF-CMA-GFS 与 WRF-GFS 表现最优，对气压下降幅度与速率的模拟最贴近观测。CMA-GFS 与 GFS 的误差较 24 h 略有上升，RMSE 分别为 11.9 与 16.7 hPa。AI 模型的误差水平与 48 h 相比变化不大，Pangu 的 RMSE 为 20.8 hPa，FuXi 为 23.1 hPa；其误差演变趋于平缓，未出现随台风强度峰值变化

而明显放大的特征，始终表现为持续的强度低估。

在 96 h 预报时效的 MSLP 预报中（图 4（h）），区域模式 WRF-GFS 与 WRF-CMA-GFS 仍然表现优异，RMSE 分别为 7.9 与 11.2 hPa，能够有效捕捉台风鼎盛阶段的极低气压特征。WRF-CMA-GFS 预报的气压最低值与观测值基本一致，仅出现时间偏早约 6 h；WRF-GFS 预报的最低气压较观测值偏高约 5 hPa。二者均能模拟出台风由缓慢发展到快速增强、再稳定维持的完整强度变化过程，且预报的 MSLP 下降速率与观测高度吻合。相比之下，AI 模型 Pangu 与 FuXi 的系统性强度低估问题尤为突出，RMSE 分别为 19.1 和 15.7 hPa，预报整体气压水平显著偏高，对台风内核极端气压场的模拟能力不足，既未捕捉到观测中的极低值，也未还原其出现时间节点，无法响应台风内核强对流抽吸导致的极端气压下降过程。其中，Pangu 预报的气压下降速率不足观测值的 1/3，在台风本应快速增强的关键时段，MSLP 仍维持在 990 hPa 附近，呈现明显的预报滞后；FuXi 同样无法刻画快速增强过程，气压变化曲线近乎平缓。



（a）（e）为 24 h 时效，（b）（f）为 48 h 时效，（c）（g）为 60 h 时效，（d）（h）为 96 h 时效。

图 5 各模式与 AI 模型近中心最大风速预报和误差

Figure 5 Forecast and error distribution of maximum near-center wind speed for various numerical models and AI model

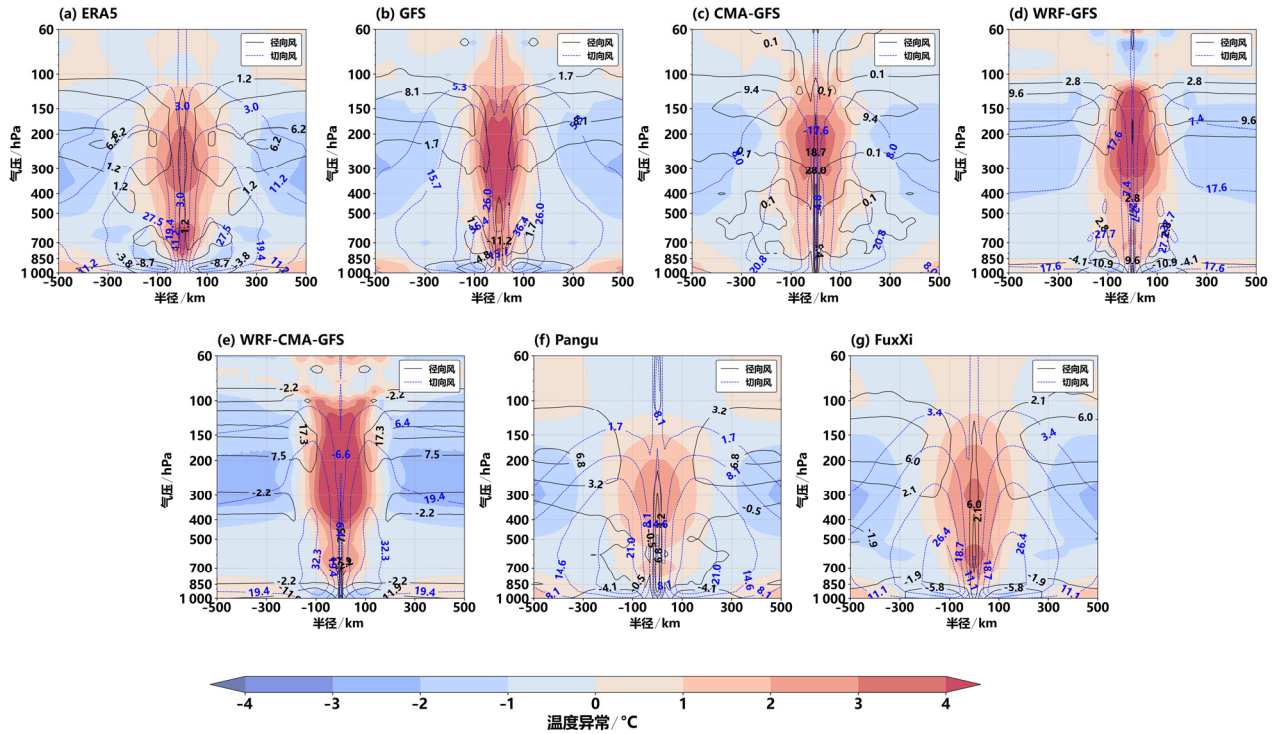
总体来看，各模式与 AI 模型强度预报误差随时效的变化特征与路径预报呈现显著差异，且 WRF 模式预报质量与初始场质量相关性较大。全球模式 GFS 呈现短期时效（24～60 h）误差较大，但误差增长较缓，在 96 h 预报时效表现最优。区域模式 WRF-GFS 与 GFS 具有相似特征，且随时效延长误差呈减小趋势，长期预报更加稳定。WRF-CMA-GFS 与 GFS 短期预报（24～60 h）表现较好，但 96 h 预报时效下误差增长较大（图 5）。受限于不同数据源在时间平均尺度上的定义差异，本文以中心最低海平面气压为台风强度评估的核心指标，近中心最大风速仅作为辅助参考。近中心最大风速预报与海平面最低气压预报在整体性能排序上一致，但区域模式 WRF-GFS 与 WRF-CMA-GFS 在风速预报中表现没有 MSLP 优异，且在 48、60 和 96 h 时效中都出现了强度预报显著偏强的情况。

2.3 物理诊断与误差成因分析

通过对台风路径、强度在各个时效预报结果的对比分析，已基本明确了不同模式与 AI 气象大模型在本次台风预报中的性能差异及误差随时间的演变特征。为进一步解释预报偏差的物理机制，本研究将从热力结构、动力结构及大尺度环流 3 个维度开展物理诊断与误差成因分析，重点考察不同模式在台风暖心、水汽输送、次级环流及副热带高压引导气流模拟上的差异，揭示路径与强度预报偏差的关键物理机制，为改进区域台风预报、优化模式初始场及提升 AI 模型物理约束提供科学依据。

2.3.1 热力结构评估 根据台风动力学经典理

论[18]，对流层中高层（200～500 hPa）暖心是台风维持与强度增强的核心热力信号，其强度、垂直伸展高度、径向集中度直接决定台风中心气压降低与最大风速发展。为此，我们选取 48 h 预报时效下 10 月 5 日 06 时各模式与 AI 模型预报的垂直热力结构剖面图（图 6）进行分析。4 种模式及 AI 模型均能再现暖心雏形，温度正异常分布在 300～700 hPa 之间，但在暖心强度、垂直连续性、径向集中度及垂直伸展高度上存在明显差异，这些结构差异与强度预报偏差在统计上具有一致性（图 6）。



（a）为 ERA5，（b）为 GFS，（c）为 CMA-GFS，（d）为 WRF-GFS，（e）为 WRF-CMA-GFS，（f）为 Pangu，（g）为 Fuxi；均为 2025 年 10 月 5 日 06 时（48 h 时效）。采用台风中心全周向方位角平均径向剖面（±500 km）。填色为温度距平，单位为（℃）；等值线为径向风与切向风，单位为（m·s^{-1}）。

图 6　各模式与 AI 模型台风垂直剖面结构

Figure 6　Vertical cross-sections of the typhoon structure for various numerical models and AI model

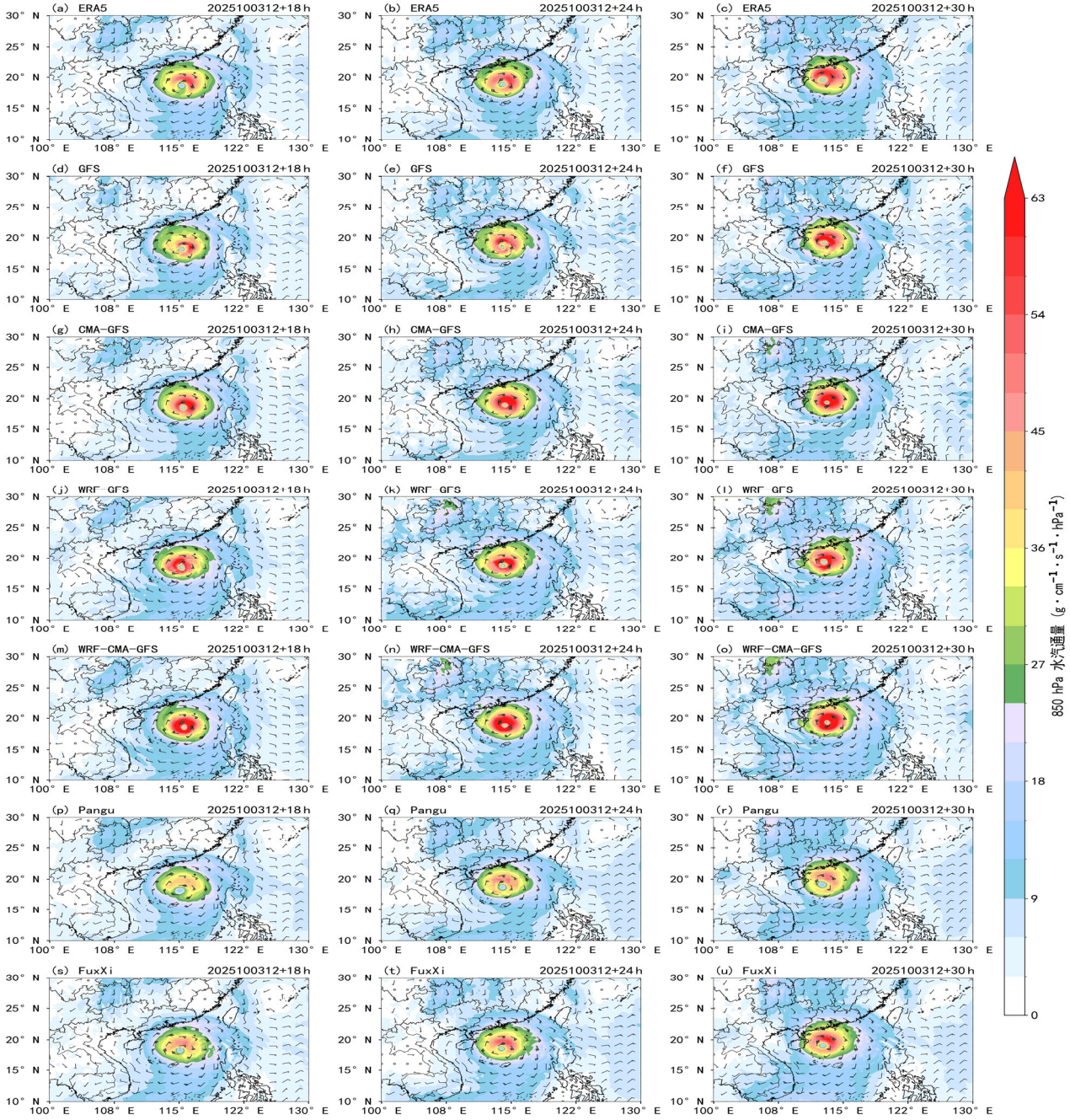
区域模式 WRF-CMA-GFS 模拟的暖异常较 ERA5 偏大，暖异常自低层向上可维持至约 100 hPa 附近，在内核半径约 100～150 km 范围内维持稳定高值区，这与该模式对台风强度预报偏强的结果相一致。但该模式能够较好地模拟出 600～700 hPa 低层大气的暖异常信号（图 6（b）），对台风结构细节的捕捉能力较强，对称结构刻画较好，体现了高分辨率模式的优越性。区域模式 WRF-GFS 对台风热力结构的模拟效果也较好，暖异常的范围与强度与 ERA5 更为接近，但同样存在暖异常偏强的情况。此外，该模式对 700 hPa 左右低层大气的暖异常信号模拟偏弱，暖核主要集中在 100～300 hPa 的高层大气。

全球模式 GFS 的中高层暖心暖异常强度偏强，暖心由 400 hPa 附近延伸至 200 hPa 且径向集中度较高（图 6（d））；其不仅中高层暖异常高估程度大于 CMA-GFS，且低层暖异常低估更明显。CMA-GFS 模拟的暖心整体偏强，中高层暖异常高于 ERA5，暖核可向上延伸至 200 hPa 附近，但低层 600～700 hPa 暖异常明显偏弱，垂直连续性一般，径向集中度较差（图 6（e））。

FuXi 模型能够模拟出暖心的大致位置，较好地模拟低层大气的暖异常信号，但其暖异常的强度与垂直深度均较 ERA5 偏弱。FuXi 在 300 hPa 附近仍有暖异常中心，但其集中半径仅在 50 km 以内，向上、下层的延伸均不足，暖异常边界更不清晰（图 6（f））。Pangu 模型的热力结构模拟偏差较 FuXi 更加突出，暖异常值仅为 2～3 ℃，高层大气的暖异常信号基本消失；径向上暖异常分布更弥散，内核半径内的高值区面积较小，暖心边界模糊（图 6（g））。

值得注意的是，FuXi 对低层暖异常的模拟及暖心径向聚合度更接近 ERA5，但强度预报偏差较大；WRF-CMA-GFS、WRF-GFS、GFS 与 CMA-GFS 的中高层暖心明显偏强，其中 GFS 径向集中度较高、垂直连续性较好，但该时刻强度预报仍然偏弱。这

一现象表明，暖心结构模拟偏差对强度预报误差的影响机制较为复杂。相关研究指出，高层暖心强度变化与台风强度变化高度一致，且高层暖心温度变化率显著大于低层，即台风强度主要由对流层中高层暖心决定[19]。因此，FuXi 中高层暖心偏弱是其强度预报偏弱的主要原因。数值模式的暖异常主要集中在中高层，强度预报偏差较 AI 模型更小，但对低空暖异常信号的捕捉能力普遍较差，低层与高层之间的对流与潜热释放不足，抽吸动力偏弱，导致该时刻台风强度较 ERA5 仍然偏弱。



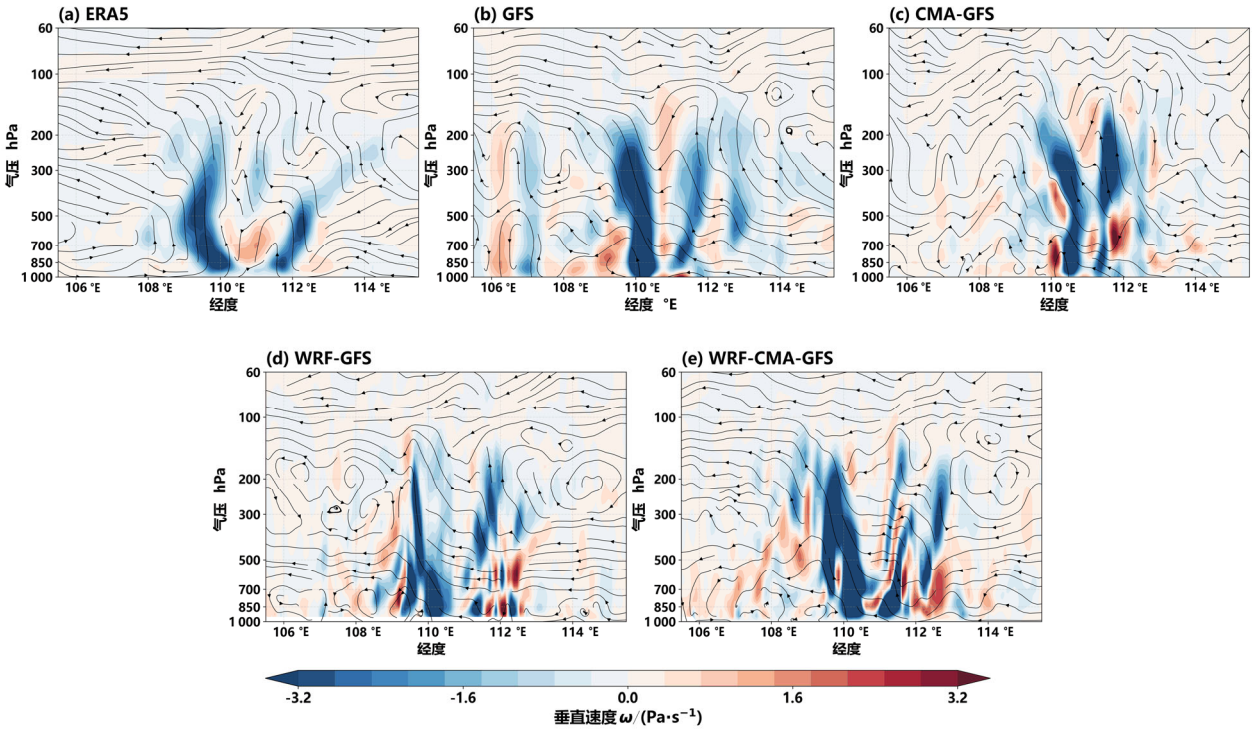
（a）（b）（c）为 ERA5，（d）（e）（f）为 GFS，（g）（h）（i）为 CMA-GFS，（j）（k）（l）为 WRF-GFS，（m）（n）（o）为 WRF-CMA-GFS，（p）（q）（r）为 Pangu，（s）（t）（u）为 FuXi；左、中、右列分别为 10 月 4 日 06 时、12 时、18 时（48 h 时效）。填色为 850 hPa 水汽通量，单位为（g·cm^{-1}·s^{-1}·hPa^{-1}）；风羽单位为（m·s^{-1}）。

图 7 各模式与 AI 模型 850 hPa 水汽通量及风场

Figure 7 850 hPa water vapor flux and wind field for various numerical models and AI model

AI 模型虽能大致捕捉台风的暖心结构，但普遍存在强度偏弱、集中性下降的问题，且误差程度与强度低估幅度呈正相关。模型间差异表现为：Pangu 额外出现高层冷异常与暖心瓦解，而 FuXi 仍保留部分中层暖心，偏差相对可控。技术原因方面：一是数据驱动的平滑效应，模型倾向于将极端暖异常

向气候平均值回归。如图 6 所示，10 月 5 日 06 时剖面图中 FuXi 和 Pangu 的温度距平标准差远小于区域模式 WRF-CMA-GFS，表明极端热力信号被压制[11,20]；二是物理机制刻画缺失，未显式融入潜热释放、暖心形成与动力反馈之间的相互作用过程，导致无法准确模拟暖心的垂直延伸与水平集中性，尤其中高层暖异常模拟存在劣势。综上，AI 模型虽能定性捕捉暖心的基本热力学结构，但因缺乏对“潜热累积—暖心增强—对流强化”正反馈循环的显式刻画，致使暖心强度、垂直延伸高度及水平集中性均出现系统性低估，且该缺陷会通过“热力异常—动力响应”链路放大强度预报偏差[18,21]。



（a）为 ERA5，（b）为 GFS，（c）为 CMA-GFS，（d）为 WRF-GFS，（e）为 WRF-CMA-GFS；均为 2025 年 10 月 5 日 06 时（48 h 时效）。流线为纬向环流，单位为（m·s^{-1}）；填色为垂直速度，单位为（Pa·s^{-1}）。

图 8　各模式台风中心垂直廓线纬向环流与垂直速度

Figure 8　Meridional circulation and vertical velocity along the typhoon central vertical profile for various numerical models

台风强度发展与暖心维持依赖低层充足的水汽供应与辐合抬升。当对流层低层水汽供应充足并伴有上升运动时，暖湿空气在上升过程中凝结释放潜热，通过 CISK 机制为热带气旋加强提供能量[22]。为进一步完善热力诊断，本文对 10 月 4 日 06 时、12 时、18 时台风增强阶段 850 hPa 水汽通量场进行对比分析。结果显示，ERA5 再分析资料中台风中心附近呈现显著的水汽通量高值区，眼墙附近水汽辐合强烈；台风的水汽主要来自南海与西太平洋的偏东气流、南海中南部的西南气流两支通道（图 7）。充沛的水汽输送为潜热释放与暖心形成提供了充足的能量支撑。

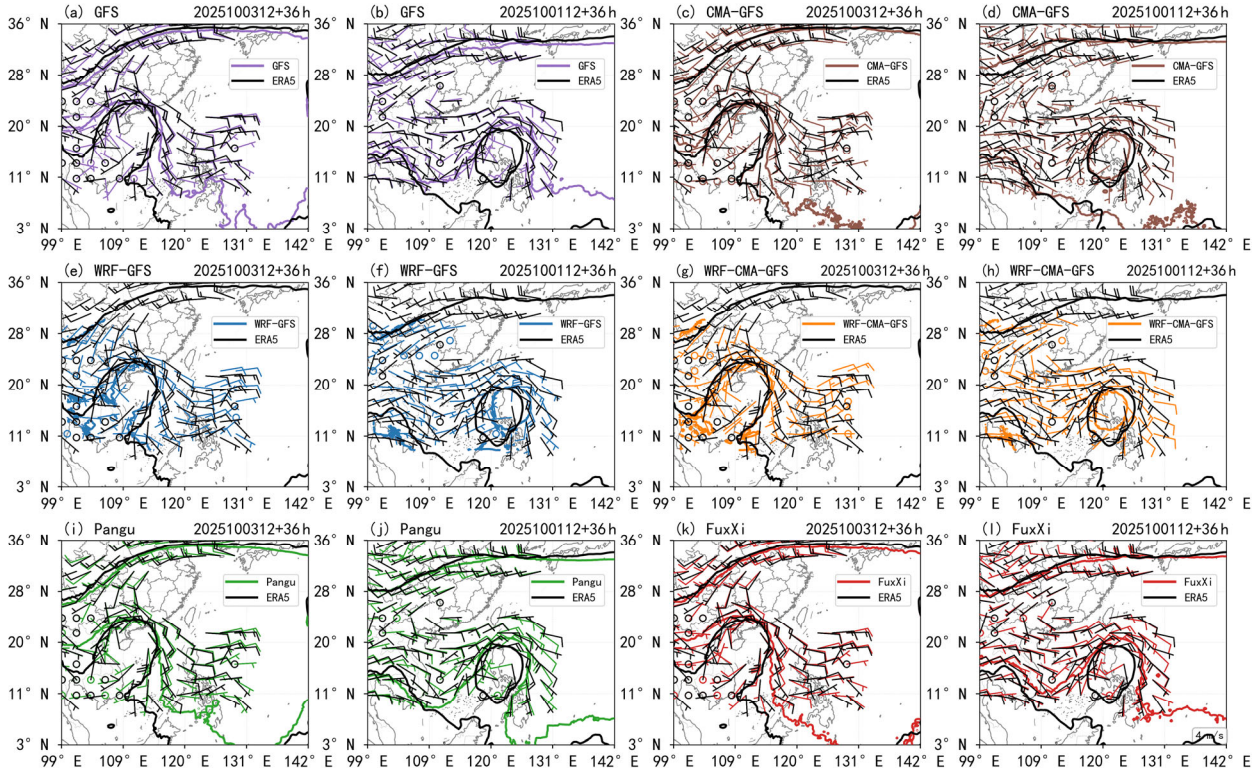
如图 7 所示，在台风增强阶段，观测中水汽通量随台风发展持续增强，辐合更为集中。区域模式 WRF-CMA-GFS 模拟的水汽通量高值区较为集中，能合理刻画增强阶段水汽输送持续加强的演变趋势，但其强度偏强，与该模式模拟的强暖心结构相对应，为台风强度维持与发展提供了稳定的水汽条件。

WRF-GFS 模拟的水汽通量高值区范围略小、强度偏弱，但对增强趋势的模拟较好。全球模式 GFS 与 CMA-GFS 能够刻画水汽输送的整体分布，但 GFS 的高值区相对分散，眼墙辐合特征不及区域模式清晰，对增强阶段水汽增幅的模拟也较为平滑。

AI 模型 Pangu 与 FuXi 的水汽通量模拟存在明显不足：水汽通量高值区强度显著偏低、分布更为平滑，台风内核区水汽辐合明显弱于观测与数值模式；在台风增强阶段，几乎无法模拟出水汽输送增长趋势，无法再现紧密的眼墙水汽输送结构。水汽供应不足、输送通道模拟偏弱与增强趋势模拟偏弱，共同导致潜热释放偏弱，直接造成 AI 模型暖心强度不足、垂直结构浅薄，与前文暖心分析结论一致，共同解释了 AI 模型强度预报系统性低估的物理成因。

各模型对水汽输送、辐合过程、水汽通道及其增强阶段演变趋势的模拟能力存在差异，这些差异通过潜热释放与热力结构形成闭环反馈，是造成台风暖心强度模拟偏差及强度预报差异的原因之一。

2.3.2 *动力结构评估* 2025 年 10 月 5 日 06 时，台风“麦德姆”经快速增强达到强度峰值，因此选取该时刻台风中心风场纬向剖面进行诊断分析。ERA5 资料显示（图 8（a）），台风具有较为清晰的眼壁区，台风中心垂直风速小于等于 1.5 Pa·s^{-1}，强烈的上升运动极值位于 850 和 300 hPa 台风眼壁附近，垂直速度小于−3 Pa·s^{-1}。



（a）（b）为 GFS，（c）（d）为 CMA-GFS，（e）（f）为 WRF-GFS，（g）（h）为 WRF-CMA-GFS，（i）（j）为 Pangu，（k）（l）为 Fuxi；其中（a）（c）（e）（g）（i）（k）为 2025 年 10 月 5 日 00 时（48 h 时效），（b）（d）（f）（h）（j）（l）为 10 月 3 日 00 时（96 h 时效）。彩色为各模式，黑色为 ERA5。等值线为 500 hPa 5 880 gpm；风羽单位为（m·s^{-1}）。

图 9 各模式与 AI 模型 500 hPa 5 880 gpm 等值线及风场

Figure 9 500 hPa 5 880 gpm contour and wind field for various numerical models and AI model

区域模式 WRF-CMA-GFS 模拟的台风眼壁纬向跨度较小，台风附近强风区深且极值较大，其对台风动力结构的模拟结果显著优于区域模式 WRF-GFS（图 8（c））。WRF-CMA-GFS 在 17°～19°N 区间呈现出宽厚的上升运动柱，强上升中心垂直速度≤−4.5 Pa·s^{-1}，自 850 hPa 连贯延伸至 200 hPa，两侧伴有对称的下沉区，构成清晰的眼墙-眼区结构。低层南风辐合、高层北风辐散的经向环流圈闭合完整，流线平滑且与垂直运动相匹配，有利于水汽向内核输送和潜热释放，为台风强度的维持与发展提供持续动力支撑。

WRF-GFS（图 8（c））无法完整预测台风的垂直结构。其上升运动柱明显偏窄且位置偏北（16°～17°N），中高层（400～200 hPa）上升运动较为活跃，强上升主要集中于 400～200 hPa 中高层，低层 850～700 hPa 东侧眼壁处上升运动微弱且断裂，垂直连贯性不足。低层辐合流场紊乱，尤为关键的是，在上升运动柱西侧约 109°E 附近出现异常下沉中心，垂直速度大于 3 Pa·s^{-1}，切断了对称的次级环流圈；低层流场紊乱并出现辐散特征，导致眼区-眼墙结构模糊，抽吸效应失效。这种垂直结构缺陷使得水汽输送与能量转换效率下降，可能是其强度预报误差（RMSE 15.6 hPa）显著大于 WRF-CMA-GFS（RMSE 6.4 hPa）（图 4（f））的重要动力原因。

全球模式 GFS 与 CMA-GFS 模拟的台风眼壁纬向跨度较 ERA5 明显偏小，垂直风速场较区域高分辨率模式更为平滑，速度峰值小于 WRF 模式，主上升区仍位于台风中心经度附近（约 109°～111°E）（图 8（d））。其中，GFS 模式对 112°～113°E 低空上升信号的模拟更好，同时在 106°～108°E 可见另

一支独立且强度较弱的上升区；该区域低层存在闭合环流，形成东西并列的双上升中心结构。图 8（e）显示，CMA-GFS 模拟的上升运动强度较 GFS 略强且范围更窄，东侧眼壁 700 hPa 附近出现异常下沉运动，且上升运动仍主要集中于中高层。

同时，WRF-CMA-GFS 的内核动力结构（图 6（a））呈现强暖心与风场紧密耦合的特征：暖心的暖异常驱动切向风峰值超过 17 $m\cdot s^{-1}$，并形成清晰的低层辐合、高层辐散结构，产生高效抽吸效应与连贯的眼墙上升运动柱；而 FuXi 与 Pangu 因暖心弱化及平滑效应，导致风场响应松散、径向风耦合失效，垂直运动支离破碎（图 6（a）（b））。这种热力-动力耦合的梯度差异，表现为 WRF-CMA-GFS 中闭合正反馈的建立与 AI 模型中的解耦失效，直接决定了强度预报精度，印证了 Emanuel[23]的暖心驱动理论及关于 AI 模型物理机制缺失的诊断。

2.3.3 *大尺度环流场模拟评估*　西太平洋副热带高压（WPSH，以下简称“副高”）是调控西北太平洋台风移动路径的核心大尺度系统，其对台风的影响主要通过引导气流实现。作为常年存在的永久性暖性深厚系统，副高在 500 hPa 等压面上表现最为典型，通常以 5 880 gpm 特征线作为其范围界定标准[1]，其外围反气旋气流构成台风移动的主导引导气流。台风作为中尺度涡旋，其移动方向与速度主要受副高外围 500 hPa 风场的走向与强度控制；该层风场被证实是调控台风路径的关键层，其气流特征直接决定台风路径的基本趋势[24]。

当副高反气旋流场呈稳定纬向分布时，引导气流持续且规律，台风路径表现出极强的可预测性；若副高的位置、强度或形态出现异常变化（如脊线北翘、西伸范围波动等），则会直接导致引导气流模拟失真。这种由副高预报偏差引发的引导气流异常，是当前台风路径预报误差的核心成因之一[24]。

基于这一机制，本研究通过对 96 h 预报时效下 2025 年 10 月 3 日 00 时、48 h 预报时效下 2025 年 10 月 5 日 00 时两个台风增强的关键时刻的 MSLP 与环流模场模拟（图 9）的可视化分析，揭示各模型路径预报误差的环流背景成因。

由图 9 可见，各模式对 WPSH 主体脊线位置的预报差异相对有限；与 ERA5 再分析场相比，不同模式对高压坝形态的刻画未表现出显著的系统性偏差，但对台风附近引导气流与 ERA5 相比却存在明显偏差。

48 h 时效下，CMA-GFS 东侧引导气流的偏北分量较大，可能导致路径北偏或北抬幅度较大；WRF-CMA-GFS 与 CMA-GFS 相似，同样受到偏北引导气流的影响导致路径偏移，而其他模式与 AI 模型未出现明显的方向偏差。

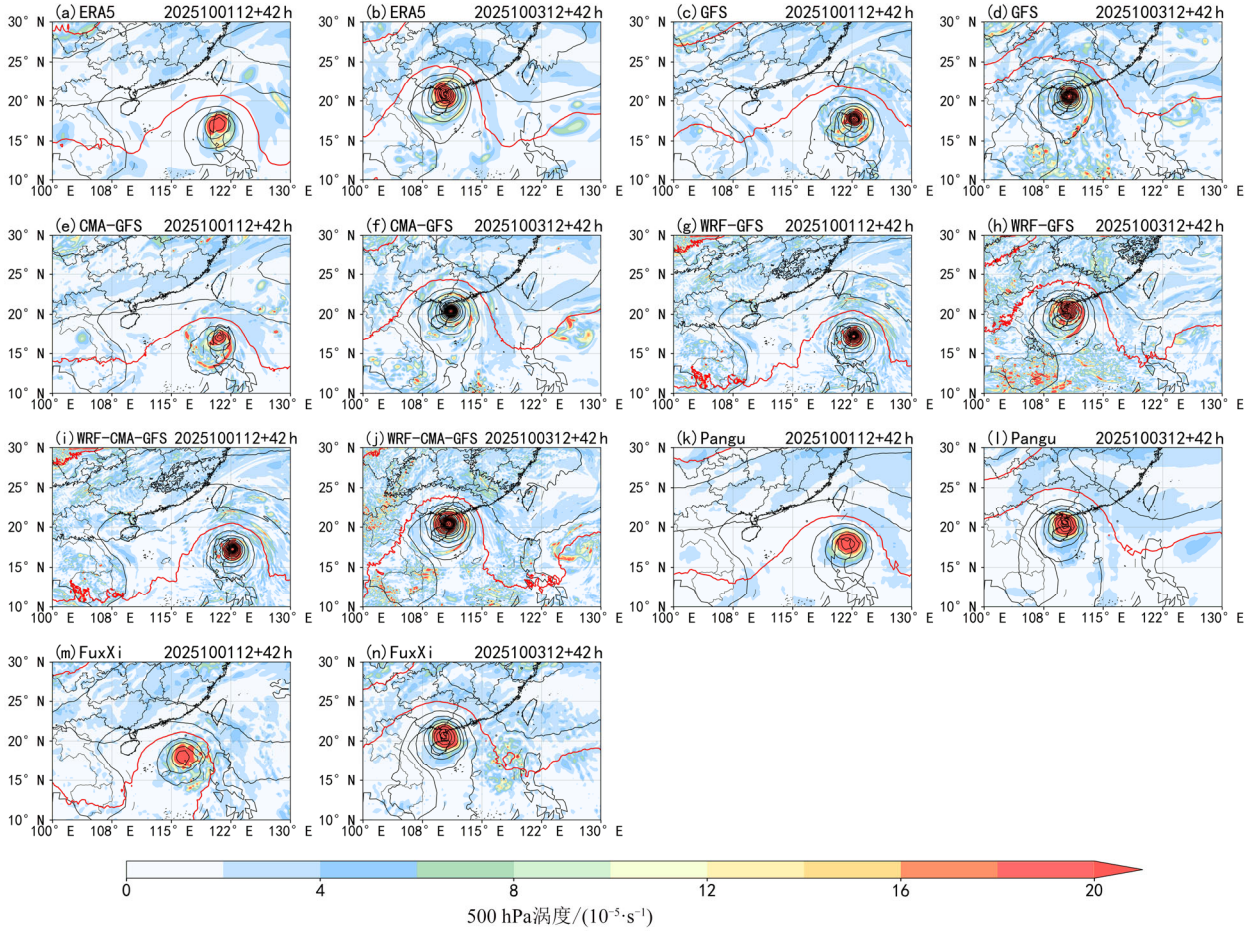
如图 9（c）所示，在 96 h 时效下，CMA-GFS 东侧引导气流偏北分量较小、偏西分量较大，导致台风北抬动力不足，进而造成 10 月 5 日 00 时路径北抬速度较慢、登陆位置偏南，这是该时效下 CMA-GFS 路径偏差的主要原因之一；WRF-CMA-GFS 也存在同样问题（图 9（k））。Pangu、FuXi、GFS 及 WRF-GFS 则与之相反，其东侧引导气流偏北分量较大、偏西分量较小，导致路径北抬迅速、登陆位置偏北等一系列问题，其中，WRF-GFS 的环流模拟效果（图 9）与 ERA5 最为接近，对应其在 96 h 预报时效下路径误差最小。

500 hPa 正涡度场是反映台风环流强度与结构完整性的关键动力特征，其分布形态与副热带高压的配置直接决定台风的引导气流与移动路径。由图 10a、b 中的 ERA5 再分析场可见，10 月 3 日 06 时，台风中心为均匀饱满的正涡度大值区，与低位势高度区重合；此时台风位于副高西南侧，涡度场形态规整、对称性良好，受稳定引导气流控制。至 10 月 5 日 06 时，台风在近海增强，正涡度中心进一步加强，涡度柱更加深厚，对应台风环流发展与暖心结构增强。

CMA-GFS 能够抓住台风与副高的相对位置，再现台风中心正涡度区与低位势区的配置关系，但如图 10（e）所示，10 月 3 日 06 时正涡度强度整体偏弱，正涡度大值区范围偏小、结构较为松散且整体对称性较差，使得台风气旋性环流被低估，引导气流强度偏弱，可能引发 48 h 时效内路径偏东、移速偏慢等预报误差。10 月 5 日 06 时涡度模拟对称性优异，结构紧凑且梯度较陡，反映出 96 h 时效下的环流模拟能力更加优异。GFS 在两时次均能较好识别台风中心位置及其闭合低位势高度结构，说明其对大尺度环流形势把握较稳定（见图 10 （c）—（d））。10 月 3 日 06 时，中心附近位势高度等值线相对偏密，梯度较强，显示出较明显的中层动力约束；但正涡度大值区形态略偏拉伸，对称性不足，提示环流质心与动力中心可能存在轻微错位。10 月 5 日 06 时，GFS 仍保持较强的位势高度梯度，但涡度分布圆整性改善幅度有限，局地非对称结构仍可能影响引导气流方向判识，进而在转向时段引入路径细节误差[25]。

WRF-CMA-GFS 在总体形态上兼具 CMA-GFS 的大尺度位置优势与 WRF 的局地细节能力。10 月

3 日 06 时，其台风中心定位较稳定，闭合环流结构较完整，正涡度核集中度优于部分原始驱动场，表明降尺度对核心结构有一定增益。10 月 5 日 06 时，环流对称性和径向集中度提升，位势高度场与涡度场匹配关系更协调，体现出在中层涡旋强度维持和结构优化方面的综合优势。



（a）（b）为 ERA5，（c）（d）为 GFS，（e）（f）为 CMA-GFS，（g）（h）为 WRF-GFS，（i）（j）为 WRF-CMA-GFS，（k）（l）为 Pangu，（m）（n）为 FuXi；（a）（c）（e）（g）（i）（k）（m）为 10 月 3 日 06 时（96 h 时效），（b）（d）（f）（h）（j）（l）（n）为 10 月 5 日 06 时（48 h 时效）。等值线为 500 hPa 位势高度，单位为（gpm）；填色为涡度，单位为（10^{-5}·s^{-1}）。

图 10 各模式与 AI 模型 500 hPa 位势高度与涡度场

Figure 10 500 hPa geopotential height and vorticity fields for various numerical models and AI model

区域模式 WRF-GFS 与 WRF-CMA-GFS 在细节尺度上呈现出更丰富的结构信息，中心区梯度刻画较清楚，其中 WRF-CMA-GFS 的对称性刻画的更好但二者的外围离散涡度斑块较多（见图 10（g）—（j）），局地噪声偏强，可能削弱大尺度引导场的一致性[26]。AI 模型模拟的涡度场大致与 ERA5 的强度相似，但位势高度等值线较为平滑，且 10 月 3 日 06 时到 10 月 5 日 06 时的涡度正值增幅较小无法反映台风近海增强过程中环流的真实发展，可能会导致引导气流强度与方向模拟失真。

3 讨论

本文对各模式及模型在台风“麦德姆”路径及强度预报中的表现进行了系统对比评估，全球模式的表现与 Sun 等[10]对超强台风“杜苏芮”的多模式评估结果基本一致，即 CMA-GFS 呈现短时效偏差较大、但长时效误差增长缓慢且稳定性突出的特征。强度预报方面，区域高分辨率模式表现显著优于全球模式，CMA-GFS、GFS 均存在明显的强度低估，对近海快速增强过程模拟能力有限 GFS 长时效易出现系统性偏北。WRF 模式与 Davis[27]及 Cha [18]等的结论一致，WRF-GFS 路径预报误差随时效延长整体增大，对台风快速增强过程的模拟能力有限，同时与 Davis 等[27]发现 WRF 存在登陆前 10 m 最大风速虚假增强的系统性偏差的结论相似，本次预报中 WRF-GFS 与 WRF-CMA-GFS 均出现了全时效虚假增强现象，但与 Cha 等[18]指出的 WRF 路径误差持续增长不同的是，本次预报中 WRF-CMA-GFS

呈现短时效偏差偏大、但长时效误差增长缓慢且稳定性更优的特点。AI 模型的台风强度预报特征与 Shi[11]和 Sahu [21]等研究结论相似，都出现了严重低估，但与其结论相悖的是，本次台风预报中 AI 模型在 96 h 预报时效下的路径预报误差不及数值模式。

本研究也存在一定局限性。首先，仅针对台风“麦德姆”单一个例开展分析，结论的普适性受限，难以推广至不同季节、不同路径与强度特征的台风。其次，区域模式未进行台风涡旋初始化与资料同化优化，初始场误差对路径模拟结果存在固有影响，无法严格验证初始场对预报性能的定量作用。最后，AI 模型仅采用确定性预报，未开展集合预报与不确定性量化，且均以 GFS 为初始场，这对路径预报产生了一定影响。

4　结论

本研究基于 CMA 最佳路径与 ERA5 再分析资料，系统对比了全球数值预报模式（GFS、CMA-GFS）、区域动力降尺度模式（WRF-GFS、WRF-CMA-GFS）与 AI 气象大模型（Pangu-Weather、FuXi）对 2025 年台风“麦德姆”在 24、48、60 与 96 h4 个预报时效的预报性能，并从路径、强度及物理结构层面进行了诊断分析。

1）路径预报具有显著的时效依赖性与模型分化特征。FuXi 在短中期时效路径精度占优，但长时效误差增长明显；Pangu 全时段偏差偏大、稳定性偏弱。GFS 短期表现良好，长时效易出现系统性偏移；CMA-GFS 短时效偏差较明显，但长时效误差增长平稳、可靠性突出。WRF-GFS 与 WRF-CMA-GFS 整体呈现短时效误差偏大、长时效稳定性更优的特点，副高及 500 hPa 引导气流模拟偏差是路径误差的主要来源。

2）强度预报随时效的演变特征与路径明显不同，且高度依赖初始场质量。全球模式与区域模式的强度误差演变规律差异显著：GFS 呈现短期误差偏大但增长平缓、96 h 表现最优的特征；WRF-GFS 与之相近，且随时效延长误差逐步收敛，长时效稳定性更佳；WRF-CMA-GFS 短期预报表现较好，但长时效误差增幅偏大。AI 模型则普遍存在强度低估问题，物理结构模拟不足是关键制约因素。

3）物理结构模拟缺陷是造成路径与强度误差的核心物理原因。AI 模型暖心偏弱、水汽辐合不足及抽吸效应失效，导致强度持续低估；WRF-CMA-GFS 的热力-动力结构模拟较为完整，对应 48 h 强度预报更优，而 WRF-GFS 则因垂直运动断裂存在明显缺陷；副高与 500 hPa 引导气流的模拟偏差直接造成各模型路径系统性偏北或偏南，是长时效路径分化的核心环流原因。

本文以近海增强台风“麦德姆”为典型个例，揭示了不同模型在不同预报时效下对台风路径与强度的预报能力差异，明确了 AI 模型强度低估、数值模式长时效稳定性不足等关键问题，并从暖心结构、水汽输送及大尺度环流等角度阐释了误差的物理机制，为业务预报中的时效策略制定和模式集成应用提供了参考。研究结果为后续 AI 模型引入动力学约束、区域模式初始化优化及多源数据一致性订正指明了改进方向，对提升我国南海及近海台风精准预报与防灾减灾能力具有应用价值。

未来可将物理约束引入 AI 模型以改善强度低估问题，并通过资料同化优化区域模式初始场；同时拓展多台风个例批量检验，进一步提升研究结论的普适性。

参考文献：

[1] Zhang G. Warming-induced contraction of tropical convection delays and reduces tropical cyclone formation[J]. Nat Commun, 2023, 14: 6274.

[2] Shan K Y, Lin Y L, Chu P S, et al. Seasonal advance of intense tropical cyclones in a warming climate[J]. Nature, 2023, 623(7985): 83-89.

[3] Liu Y H, Guan S D, Lin I I, et al. Effect of storm size on sea surface cooling and tropical cyclone intensification in the western north Pacific[J]. J Clim, 2023, 36(20): 7277-7296.

[4] Saha S, Moorthi S, Wu X R, et al. The NCEP climate forecast system version 2[J]. J Clim, 2014, 27(6): 2185-2208.

[5] 沈学顺, 苏勇, 李兴良, 等. 中国业务数值天气预报自主发展中的理论和方法创新[J]. 气象, 2025, 51(11): 1293-1320.

[6] 张进, 孙健, 沈学顺, 等. CMA-GFS V4.0 模式关键技术研发和业务化[J]. 应用气象学报, 2023, 34(5): 513-526.

[7] 韩威, 尹若莹, 李俊, 等. 静止气象卫星大气廓线目标观测及在高影响天气预报中的影响[J]. 中国科学: 地球科学, 2025, 55(4): 991-1005.

[8] Bi K F, Xie L X, Zhang H H, et al. Accurate medium-range global weather forecasting with 3D neural networks[J]. Nature, 2023, 619(7970): 533-538.

[9] Chen L, Zhong X H, Zhang F, et al. FuXi: a cascade machine learning forecasting system for 15-day global weather forecast[J]. npj Clim Atmos Sci, 2023, 6: 190.

[10] Sun X Y, Zhong X H, Xu X Z, et al. A data-to-forecast machine learning system for global weather[J]. Nat Commun, 2025, 16: 6658.

[11] Shi Y, Hu R, Wu N G, et al. Comparison of AI and NWP models in operational severe weather forecasting: a study on tropical cyclone predictions[J]. J Geophys Res Mach Learn Comput, 2025, 2(2): e2024JH000481.

[12] Bonavita M. On some limitations of current machine learning weather prediction models[J]. Geophys Res Lett, 2024, 51(12): e2023GL107377.
[13] 中华人民共和国应急管理部. 国家防灾减灾救灾委员会办公室 应急管理部发布2025年10月全国自然灾害情况[EB/OL]. (2025-11-11)[2026-03-12]. https://www. mem. gov.cn/xw/yjglbgzdt/202511/t20251111_585250. shtml.
[14] Ying M, Zhang W, Yu H, et al. An overview of the China meteorological administration tropical cyclone database[J]. J Atmos Ocean Technol, 2014, 31(2): 287-301.
[15] Skamarock W C, Klemp J B, Dudhia J, et al. A description of the advanced research WRF Model version 4[R]. NCAR tech. note ncar/tn-556+ str, 2019, 145(10.5065).
[16] Lam R, Sanchez-gonzalez A, Willson M, et al. Learning skillful medium-range global weather forecasting[J]. Science, 2023, 382(6677): 1416-1421.
[17] Yamaguchi M, Ishida J, Sato H, et al. WGNE intercomparison of tropical cyclone forecasts by operational NWP models: a quarter century and beyond[J]. Bull Am Meteor Soc, 2017, 98(11): 2337-2349.
[18] Cha D H, Wang Y Q. A dynamical initialization scheme for real-time forecasts of tropical cyclones using the WRF model[J]. Mon Weather Rev, 2013, 141(3): 964-986.
[19] Zhang D L, Chen H. Importance of the upper-level warm core in the rapid intensification of a tropical cyclone[J]. Geophys Res Lett, 2012, 39(2): 2806.
[20] Charlton-perez A J, Dacre H F, Driscoll S, et al. Do AI models produce better weather forecasts than physics-based models A quantitative evaluation case study of Storm Ciarán[J]. npj Clim Atmos Sci, 2024, 7: 93.
[21] Sahu P L, Sandeep S, Kodamana H. Evaluating global machine learning models for tropical cyclone dynamics and thermodynamics[J]. J Geophys Res Mach Learn Comput, 2025, 2(2): e2025JH000594.
[22] 高拴柱, 吕心艳, 王海平, 等. 热带气旋莫兰蒂(1010)强度的观测研究和增强条件的诊断分析[J]. 气象, 2012, 38(7): 834-840.
[23] EMANUEL K A. An air-sea interaction theory for tropical cyclones. part I: steady-state maintenance[J]. J Atmos Sci, 1986, 43(6): 585-605.
[24] 王志烈, 胡坚, 丁一汇. 西北太平洋风场对台风路径的影响[J]. 应用气象学报, 1991, 2(4): 362-368.
[25] 吴胜蓝, 漆梁波, 余晖. 登陆台风热力不对称结构对路径偏折的影响[J]. 气象, 2025, 51(6): 675-685.
[26] 陈联寿. 热带气旋研究和业务预报技术的发展[J]. 应用气象学报, 2006, 17(6): 672-681.
[27] Davis C, Wang W, Chen S S, et al. Prediction of landfalling hurricanes with the advanced hurricane WRF model[J]. Mon Weather Rev, 2008, 136(6): 1990-2005.

本刊青年编委杨天元副教授在 *Horticulture Research* 刊发关于茶树品质调控的研究成果

安徽农业大学茶业学院杨天元副教授团队与云南农业大学农学与生物技术学院赵明教授团队在国际期刊 *Horticulture Research*（IF=8.5）上发表了题为“LncR12304.1-miR1507c-*CsNADH-GOGAT* ceRNA module regulates amino acid biosynthesis in tea plant (*Camellia sinensis*)”的研究论文。

竞争性内源性 RNA（competitive endogenous RNA，ceRNA）可通过微小 RNA 响应元件竞争性结合 miRNA，调控靶基因 mRNA 表达。谷氨酸合成酶（glutamate synthase，GOGAT）作为高等植物谷氨酰胺合成酶/谷氨酸合成酶（GS/GOGAT）循环的限速酶，是氮代谢与氨基酸合成的关键酶，分铁氧还蛋白型（Fd 型，ferritin type）和 NADH 还原型，二者分别主要负责植物叶片和根系中的氨同化过程。GOGAT 与氮素利用效率及碳氮平衡密切相关。但因 GOGAT 序列较长，功能研究较为匮乏，且转录后水平非编码 RNA 及 ceRNA 网络对其的靶向调控报道极少。在茶树中，氨基酸是核心风味与功能成分，L-茶氨酸为其特征氨基酸，目前虽鉴定出部分调控氨基酸合成的转录因子及转运蛋白，但非编码 RNA 的层级调控研究不足，已报道的茶树 ceRNA 网络多与抗逆、白化相关，尚未见其调控氨基酸代谢的报道。

研究首次在茶树中鉴定并验证了 lncR12304.1-miR1507c-*CsNADH-GOGAT* ceRNA 调控模块，明确了其在氨基酸合成中的作用机制，填补了茶树氨基酸代谢 ceRNA 网络研究的空白。研究拓展了 miR1507 家族的功能范围，证实其不仅参与豆科植物抗逆反应，还可调控茶树氨基酸代谢，同时揭示了 lncR12304.1 作为 ceRNA 的分子海绵功能，丰富了植物氮代谢转录后调控的理论体系，为茶树氨基酸代谢的分子调控机制提供了直接实验证据。该成果为茶树高氨基酸品种的靶向育种与改良提供了科学理论依据，推动了茶树次生代谢 ceRNA 机制的研究，对提升茶叶品质具有重要应用价值。

安徽农业大学茶业学院杨天元副教授与云南农业大学农学与生物技术学院赵明教授为论文共同通讯作者，该研究得到云南省基础研究计划项目（202201AT070253、202501AS070039）、茶树种质创新与资源利用全国重点实验室开放基金（NKLTOF20240110）及安徽省自然科学基金（2408085MC077）等项目的资助。

（内容来源于创新生物学公众号）