



FourCastNet 3天氣尺度系集預報 與初始場敏感度評估

陳郁涵¹ 連國淵¹ 蘇胤瑞² 沙聖浩¹ 劉正欽¹ 郭珮萱¹ 李尚恩¹
李志昕¹ 蕭玲鳳¹

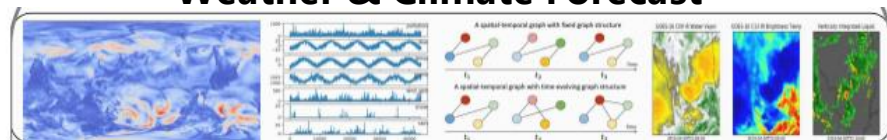
¹中央氣象署科技發展組 ²中央氣象署海象氣候組

Multiple MLWP Models Announced

Worldwide (before Sep. of 2023)

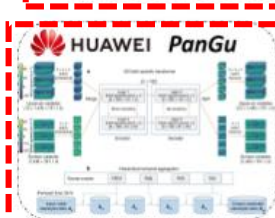
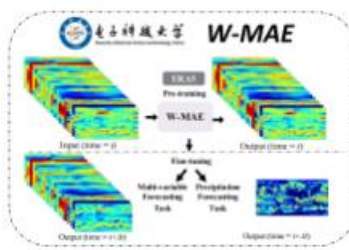
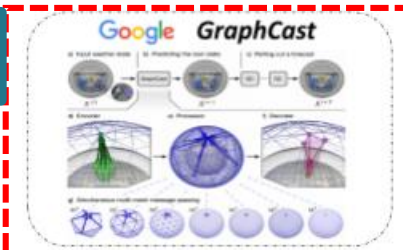
Integrate large-scale models currently used in weather and climate forecasting, such as Pangu, FuXi, GraphCast, etc.

Weather & Climate Forecast

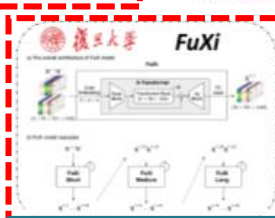


FengWu
by Shanghai
AI
Laboratory

GraphCast
by Google



PanGu-Weather
by HUAWEI



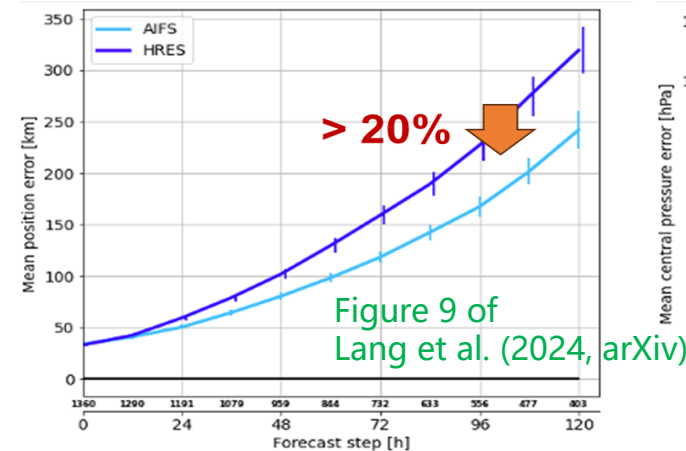
FuXi by
Fudan
University



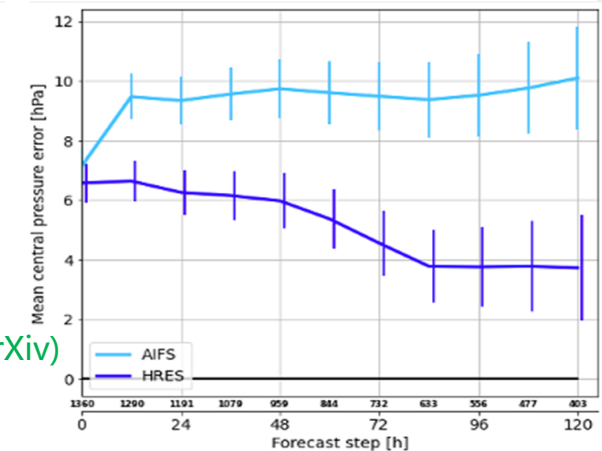
FourCastNet
by NVIDIA

ECMWF launched "AIFS" (start around Oct. 2023)

Track forecast errors



Intensity forecast errors



- ECMEF AIFS' tropical cyclone track forecast **outperforms their IFS by >20% !!** (two-year global average)

AIFS OPERATIONAL

Begins → 25 February 2025 06 UTC

★ AIFS v2 Release On 12 May 2026

CWA Multi-model AI Ensemble: Application and Verification for Typhoon Danas

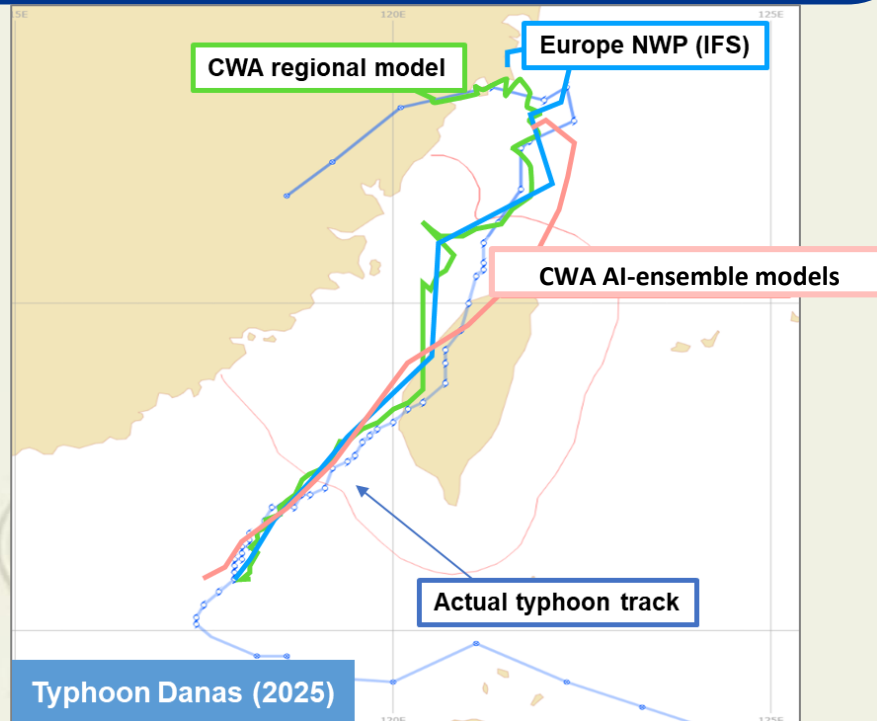
18 Global AI models
→ **Multi-Model AI Ensemble**

6 Open-source models
Including Aurora, FCN2, GraphCast, FuXi,
FengWu, Pangu

3 Initial data
USA NWP (NCEPGFS)
Europe NWP (IFS)
CWA Global NWP (TGFS)

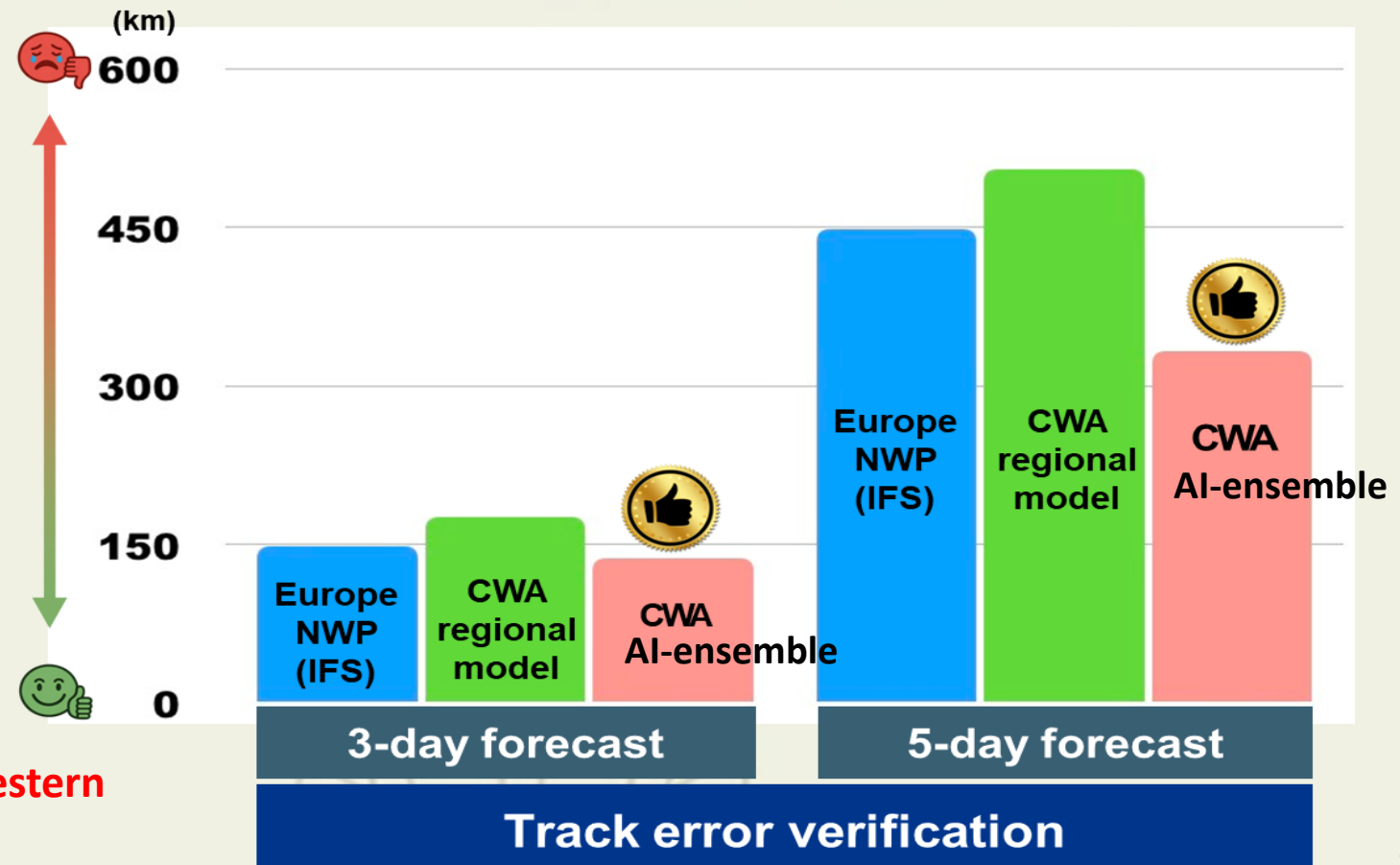


Forecast tracks using AI/ML and
physics-based models (20 UTC, July 5)

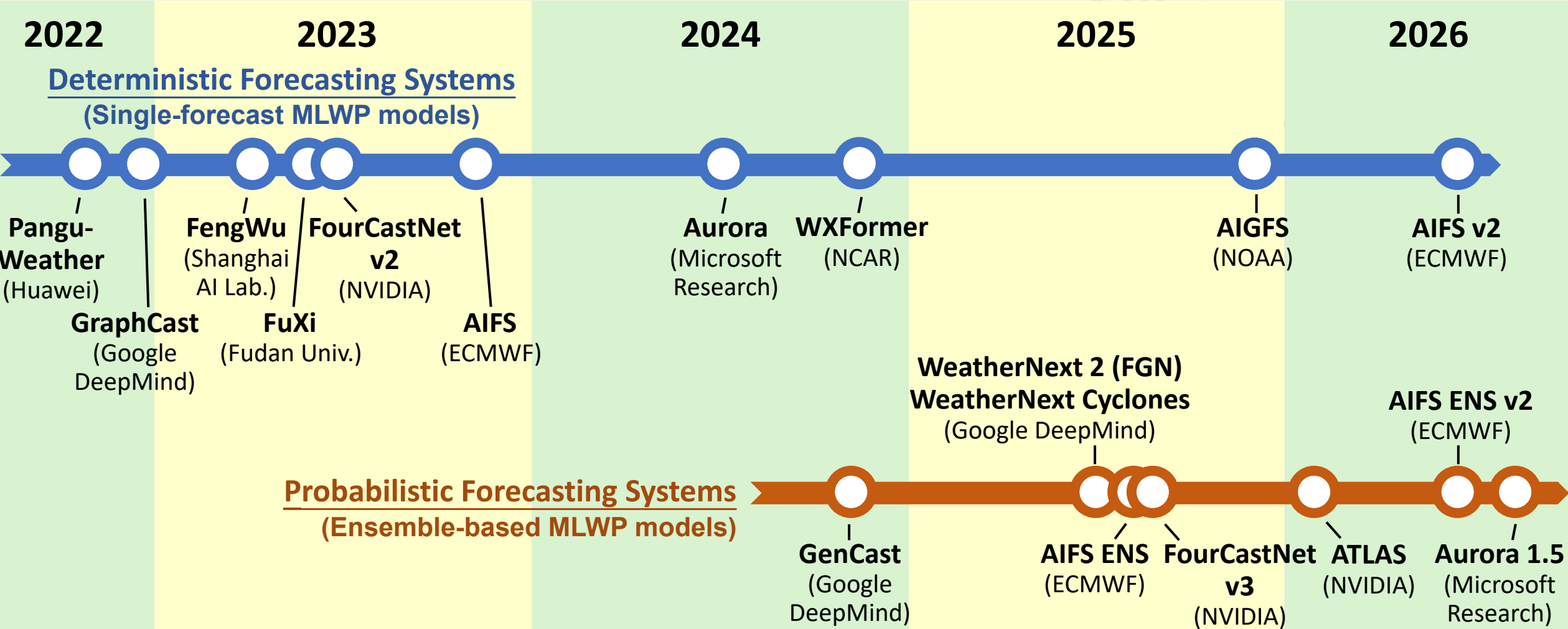


Fewer than five typhoons made landfall in southwestern Taiwan in the ~40-year training data.

2024-2025 Northwest Pacific typhoon cases



Evolution of Global AI/ML Weather Prediction (MLWP) Models



⌘ Operational MLWP forecasts currently depend entirely on conventional NWP models and data assimilation systems for **training data** and **initial forecast conditions**.

From Deterministic to Probabilistic AI Weather Forecasting

➤ Deterministic forecasting systems

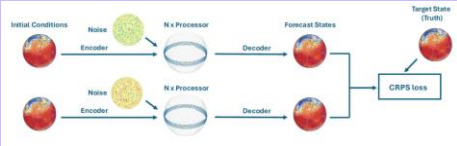
- Representative models: GraphCast, Pangu-Weather, FourCastNet
- Generate a single deterministic forecast for each initial condition
- Optimized with deterministic metrics (e.g., RMSE)
- No uncertainty representation
- Even with multiple initial conditions, ensembles tend to be under-dispersed
- Overly smooth forecasts at longer lead times

➤ Probabilistic forecasting systems

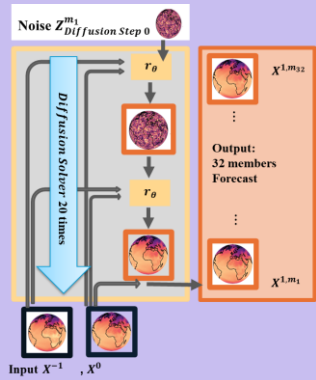
- Representative models: GenCast, AIFS ENS, FourCastNet3, WeatherNext 2
- Native ensemble generation through stochastic perturbations
- Trained with probabilistic metrics (e.g., CRPS)
- Capable of uncertainty quantification
- Better maintain weather-system structure and spectral characteristics at longer lead times
- Show potential for subseasonal forecasting

Evaluation of Global Ensemble MLWP

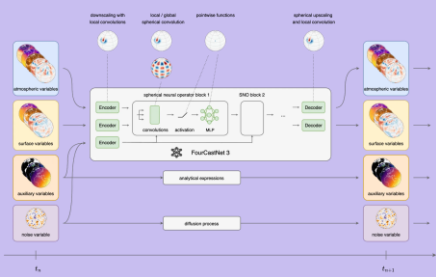
Candidate AI-based ensemble model



ECMWF
AIFS ENS



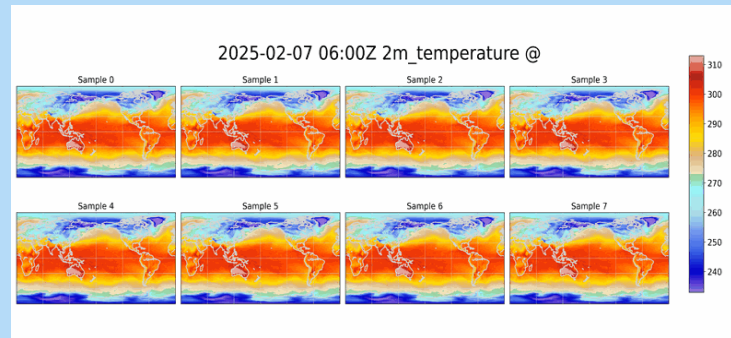
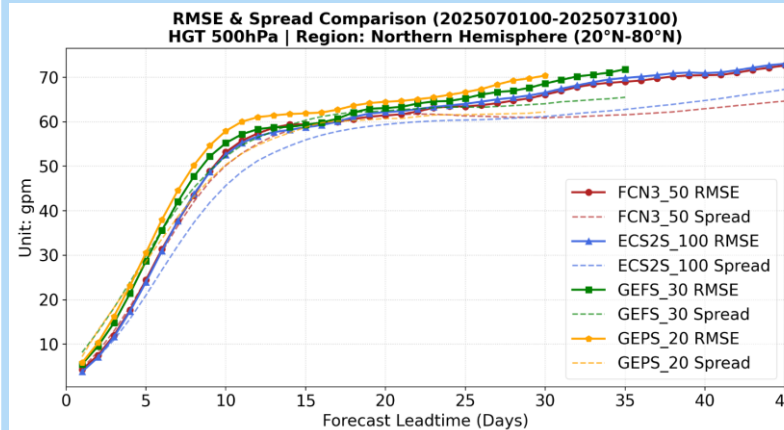
Google
DeepMind
GenCast



NVIDIA
FourCastNet3
(FCN3)

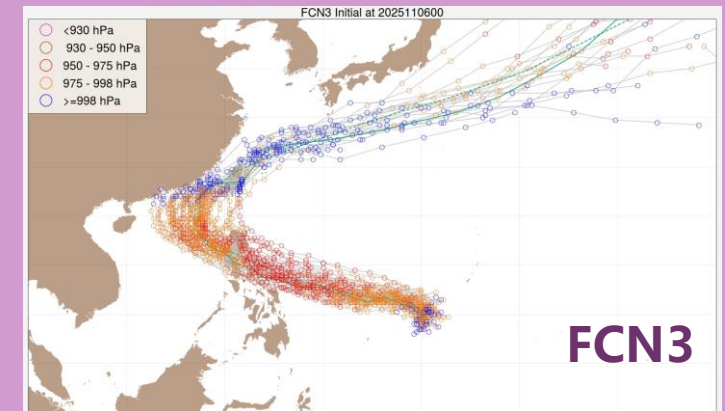
Evaluation and assessment

Statistical evaluation



Applications

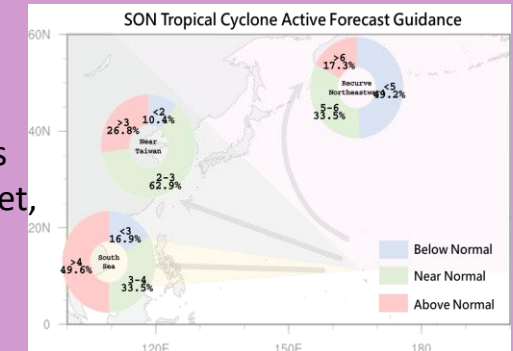
Typhoon track ensembles



FCN3

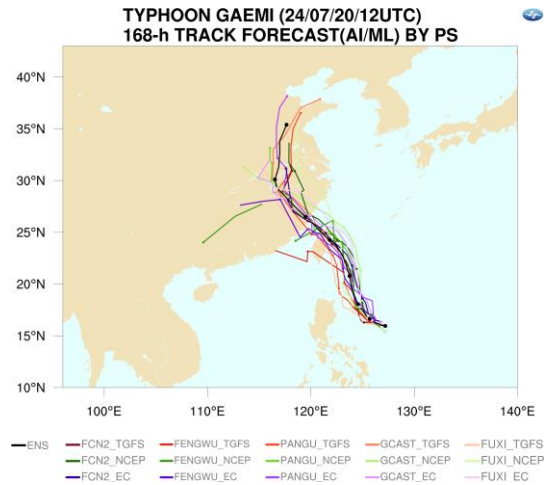
Subseasonal application

To Assess the MLWP on the subseasonal forecast in terms of monsoon onset, TC forecast and MJO

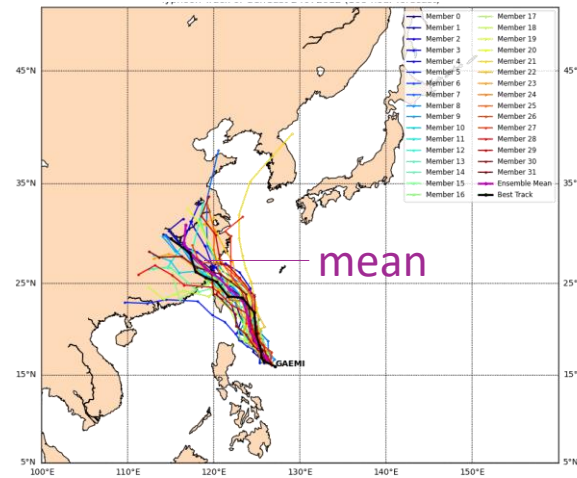


GenCast Ensemble Typhoon Track Forecasts

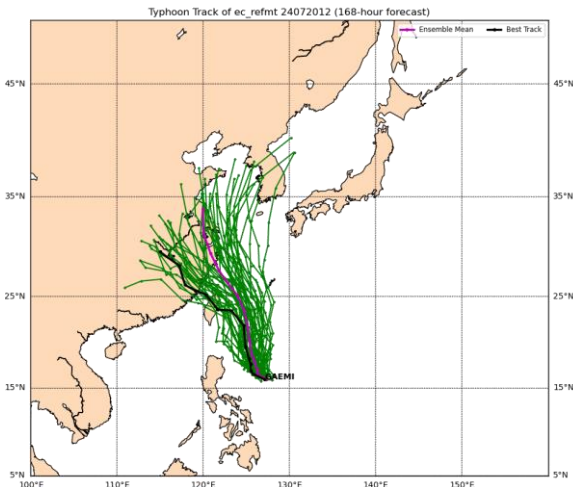
CWA AI-ensemble



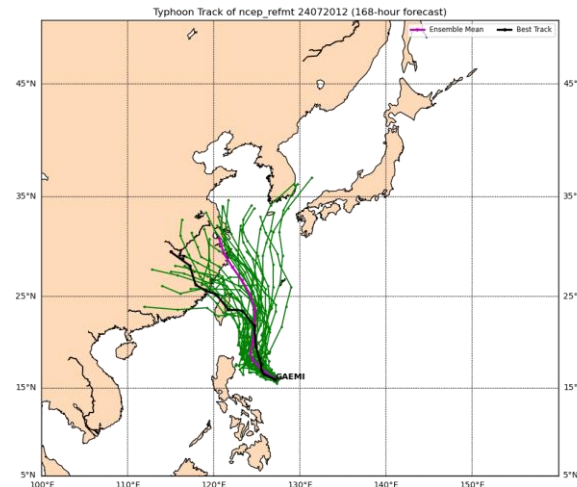
GenCast



IFS Ensemble (Physics-based)



GEFS Ensemble (Physics-based)



Typhoon Gaemi (2024)

Initial time: 2024/07/20 12Z

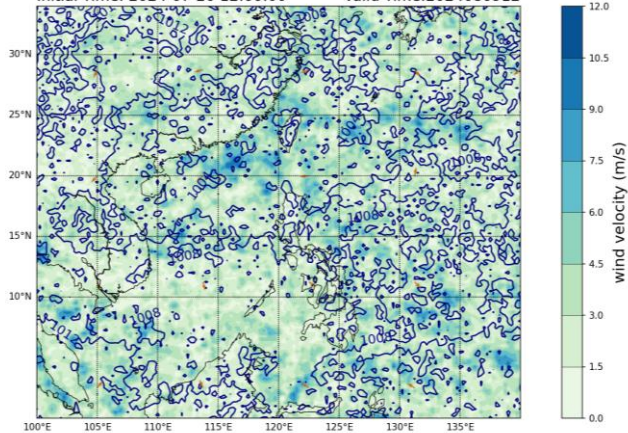
- **CWA AI-ensemble vs. GenCast:**
 - GenCast shows a larger spread in typhoon track forecasts.
 - **GenCast vs. traditional physics-based ensembles:**
 - Track spread is comparable.
 - In this case, GenCast's ensemble mean track is more accurate, and similar to other deterministic MLWP ensemble means.
- **GenCast is relatively slow on CWA's GPU environment, limiting its current operational applicability.**

Evaluation of GenCast, AIFS ENS, and FCN3 Forecast Performance

GenCast

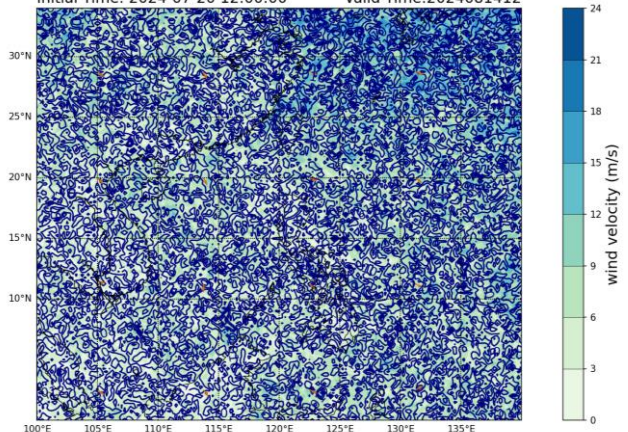
20 days

10 m wind velocity and direction, mean sea level pressure
Initial Time: 2024-07-20 12:00:00 Valid Time:2024080912



25 days

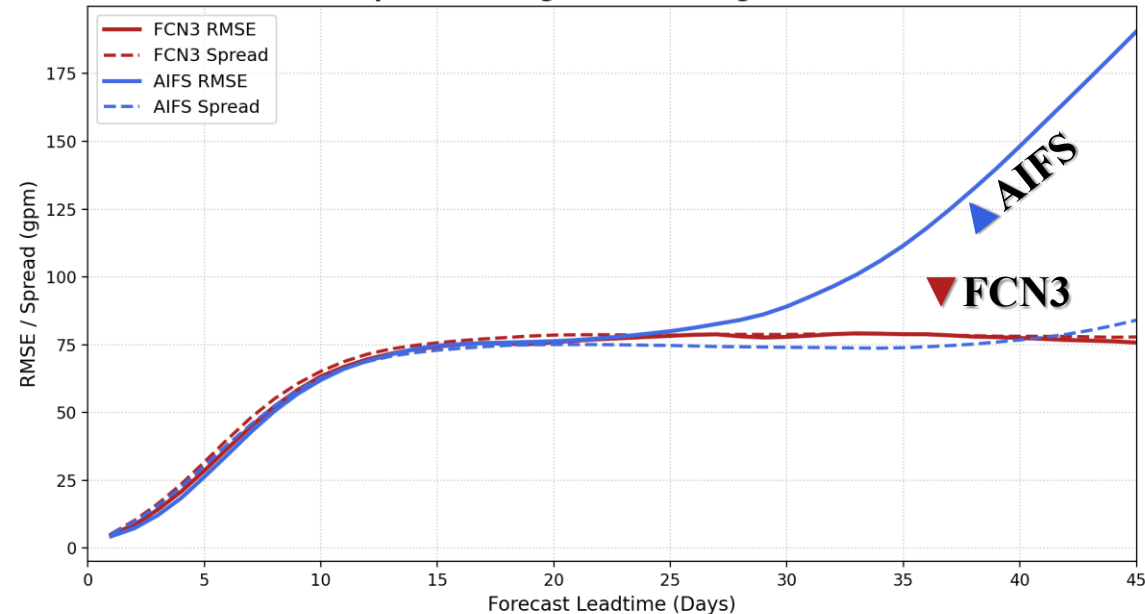
10 m wind velocity and direction, mean sea level pressure
Initial Time: 2024-07-20 12:00:00 Valid Time:2024081412



(Image courtesy of 凌文海)

RMSE & spread

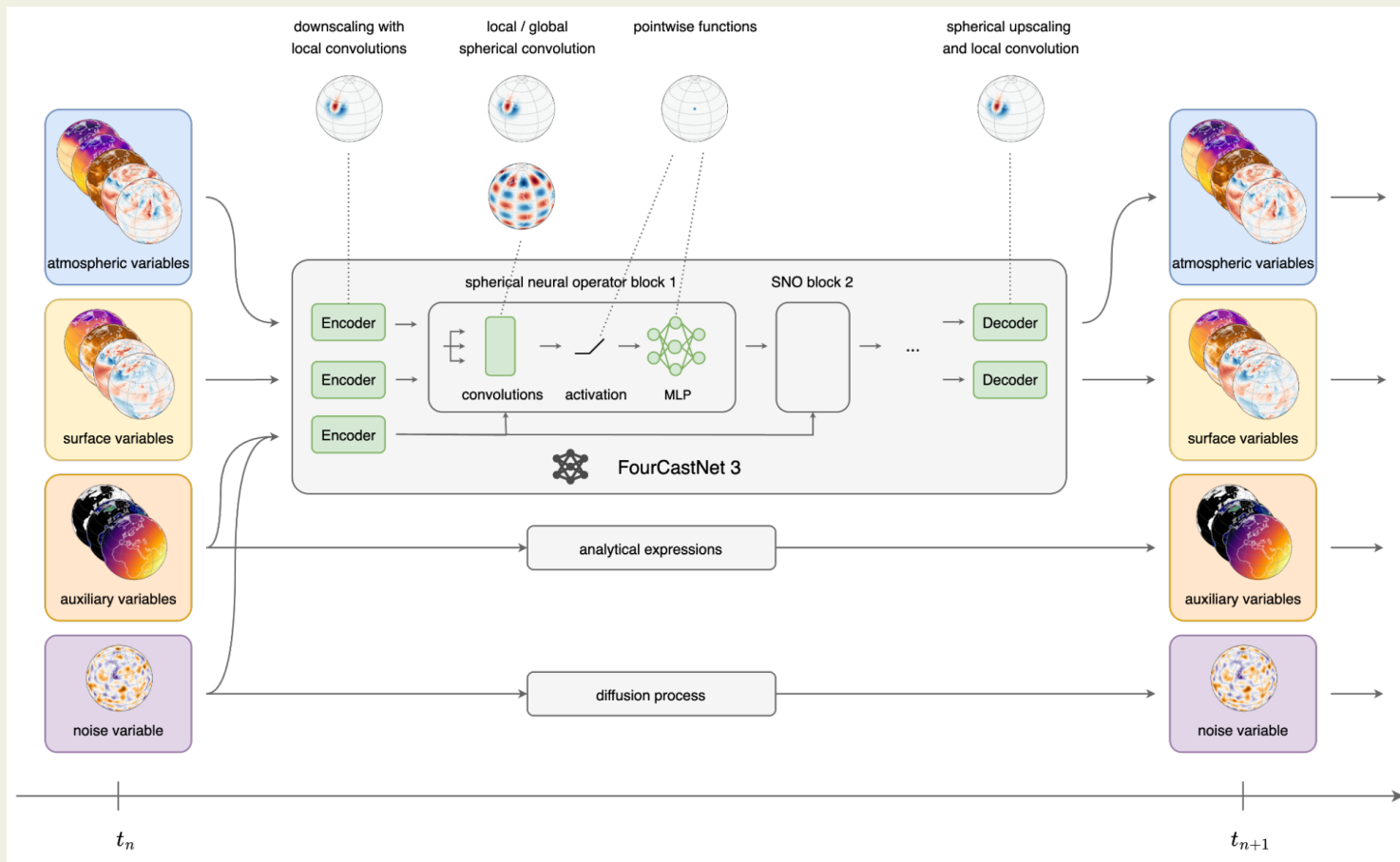
RMSE & Spread Comparison (2025051500-2025061500)
Geopotential Height 500hPa (Region: Global)



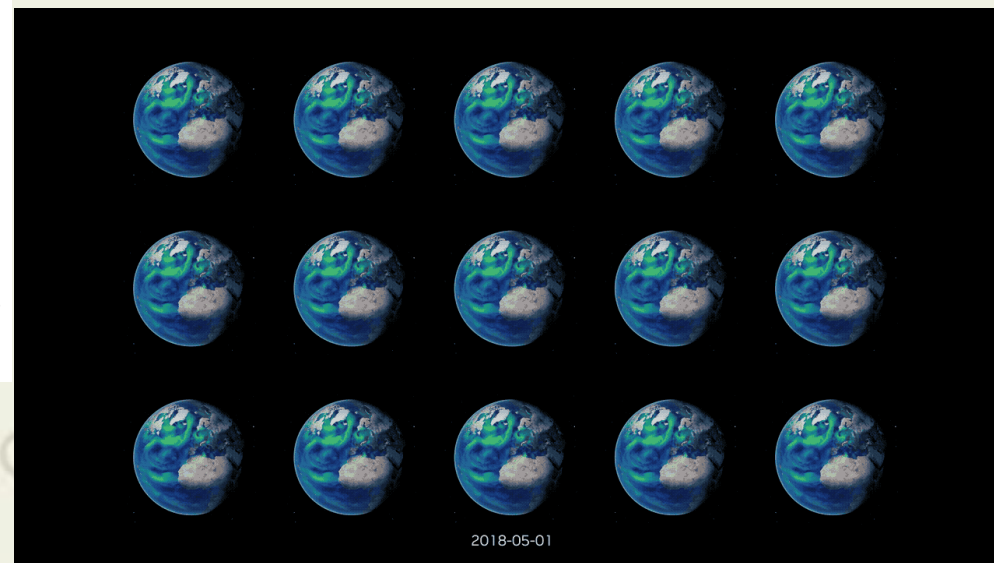
- **GenCast:**
Small-scale noise emerges at longer lead times (> 15 days).
- **AIFS ENS:**
Unrealistic error growth at longer lead times.
- **FCN3:**
More reasonable error growth through Day 45, indicating the greatest potential for extended-range forecasting.

FourCastNet v3 (FCN3)

➤ FCN3 architecture



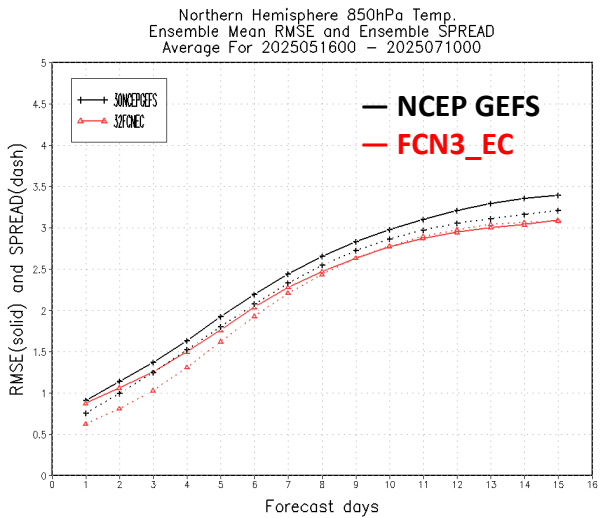
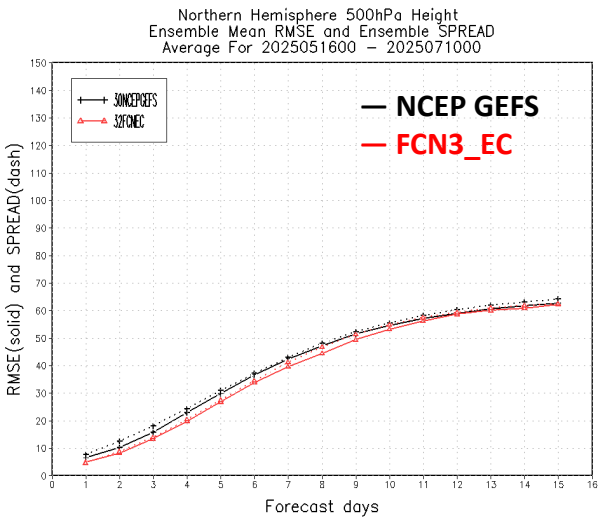
➤ **2-week rollout of 15 FourCastNet3 ensemble members, displaying surface wind speeds during this period.**



<https://developer.nvidia.com/blog/fourcastnet-3-enables-fast-and-accurate-large-ensemble-weather-forecasting-with-scalable-geometric-ml/>

Ensemble Forecast Verification: FCN3 (MLWP) vs. NCEP GEFS (NWP)

- Initialization Period: 15 May – 10 July 2025 (00Z)
- Verification Lead Time: 360 h (15 days)
- Ensemble Members:
FCN3EC (32), NCEPGEFS (30)
- Verification Reference: Respective analysis fields



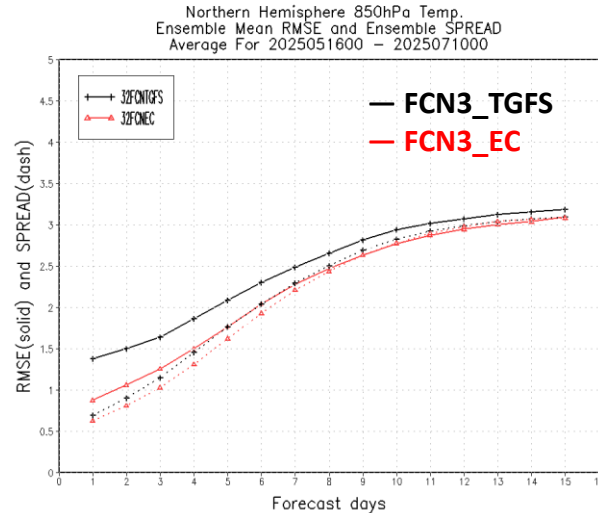
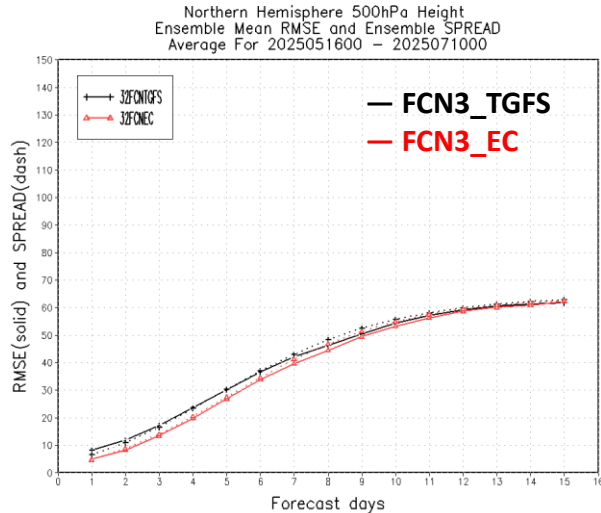
FCN3EC **better/worse** than
NCEP GEFS

EMC Verification Scorecard	
Symbol Legend	
▲	32FCNEC is better than 32FCNTGFSat the 99.9% significance level
△	32FCNEC is better than 32FCNTGFS at the 99% significance level
△	32FCNEC is better than 32FCNTGFS at the 95% significance level
○	No statistically significant difference between 32FCNEC and 32FCNTGFS
▼	32FCNEC is worse than 32FCNTGFS at the 95% significance level
▽	32FCNEC is worse than 32FCNTGFS at the 99% significance level
▽	32FCNEC is worse than 32FCNTGFS at the 99.9% significance level
	Not statistically relevant
Start Date: 2025051600	
End Date: 2025071000	

		Northern Hemisphere							Southern Hemisphere							Tropical						
		day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 1	day 2	day 3	day 4	day 5	day 6	day 7
Anomaly Correlation	Z1000	▲	△	△	○	△	▲	▲	▲	▲	▲	▲	▲	▲	○	○	△	○	○	△	▲	▲
	Z500	▲	▲	△	△	▲	△	▲	▲	▲	▲	△	▼	▲	▲	△	▲	▲	▲	▲	▲	▲
	T850	▼	▲	▲	▲	▲	○	○	○	▼	○	▲	▲	▲	○	○	△	▲	▲	△	▲	▼
	T2M	▼	▼	▼	▼	▼	▼	▼	▼	▼	○	▲	▼	▲	▲	▼	▼	▼	▼	▼	▼	▼
	U10M	▼	▼	○	○	△	△	▼	▼	▼	▲	▼	○	○	○	▼	▼	○	○	△	▲	▲
	V10M	▼	▼	○	○	○	▲	▼	▼	▼	△	○	▼	▲	▲	▼	▼	M	▼	▼	△	○
	U850	▲	▲	▲	▲	△	▲	▲	○	▼	○	△	▼	△	▼	▲	△	○	▲	▲	▲	▲
	V850	○	○	○	△	△	▲	▲	○	○	○	○	○	○	▲	○	○	△	△	▲	▲	△
	U250	▲	▲	▲	○	○	○	▲	○	▲	▲	▲	▼	▲	○	○	○	○	○	▲	▲	▲
	V250	▲	▲	▲	△	△	○	▲	△	▲	▲	▲	▲	▲	○	○	○	▼	○	△	▲	▲
RMSE	PPRMSL	○	○	△	○	△	▲	▲	▲	▲	▲	▲	▲	▲	▲	○	○	○	○	△	▲	▲
	Z1000	▲	△	▲	▲	▲	▲	▲	▲	▲	▲	▲	○	▲	▲	▼	▼	▼	▼	▼	△	▲
	Z500	▲	▲	△	○	▲	○	▲	▲	▲	▲	▲	○	▲	▲	○	▲	▲	▲	▲	▲	▲
	T850	▼	○	△	△	▲	△	▼	▼	▼	▼	▼	▼	△	▲	▼	▼	▼	▲	○	▼	▼
	T2M	▼	▼	▼	▼	▼	▼	▼	○	○	○	▲	○	△	▲	▼	▼	▼	▼	▼	▼	▼
	U10M	▼	○	△	△	△	▲	▲	▼	▼	▼	▲	▼	△	○	▼	▼	▼	▼	▼	▼	▲
	V10M	▼	○	△	△	○	▲	▲	▼	▼	△	▲	△	▲	▲	▼	▼	▼	▼	○	○	○
	U850	▲	▲	▲	△	▲	▲	▲	▼	▼	○	▲	▼	▲	▲	M	▼	▼	▼	▼	△	▲
	V850	▼	○	○	▲	▲	○	▲	○	○	○	○	▼	▲	▲	▼	▼	M	▼	○	▼	▼
	U250	▲	▲	▲	△	○	○	▲	○	▲	▲	▲	○	▲	▲	△	○	○	▼	○	▲	▲
	V250	▲	▲	▲	▲	○	○	▲	△	▲	▲	▲	▲	▲	○	▼	○	▼	○	△	▲	▲
	PPRMSL	○	○	▲	▲	△	▲	▲	▲	▲	▲	▲	△	▲	▲	▼	▼	▼	▼	▼	▼	▲

FCN3 Ensemble Forecasts: TGFS vs. EC Initial Conditions

- **Experiment (Initial Conditions):** FCN3EC (EC) vs. FCN3TGFS (TGFS)
- **Initialization Period:** 15 May – 10 July 2025 (00Z)
- **Verification Lead Time:** 360 h (15 days)
- **Ensemble Members:**
FCN3EC (32), FCN3TGFS (32)
- **Verification Reference:** Respective analysis fields



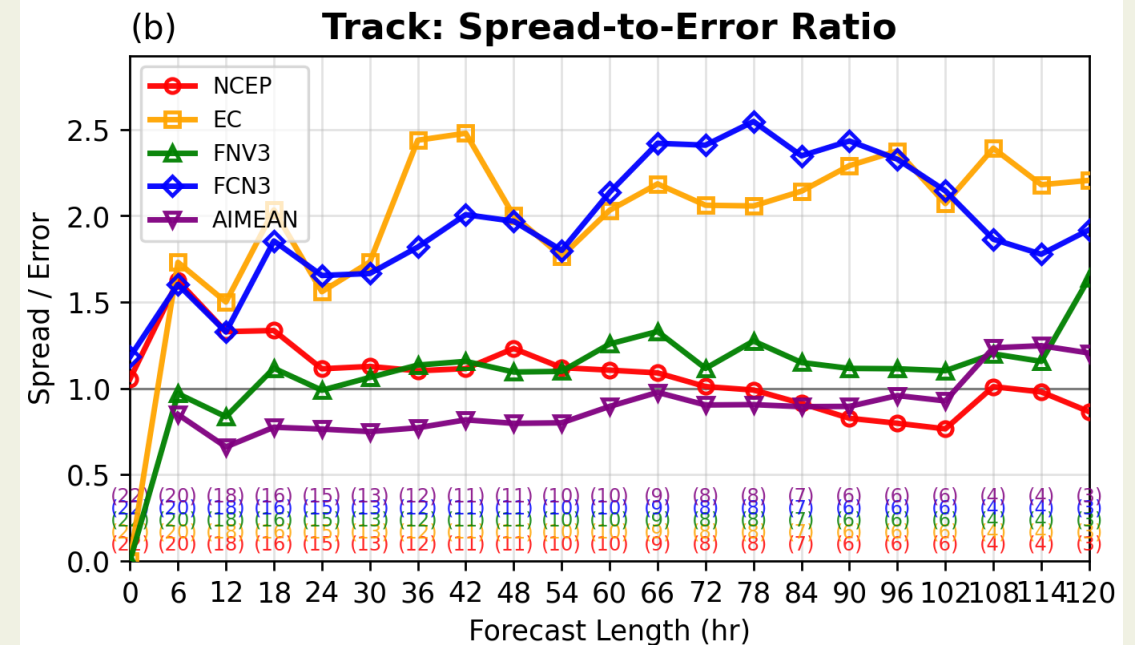
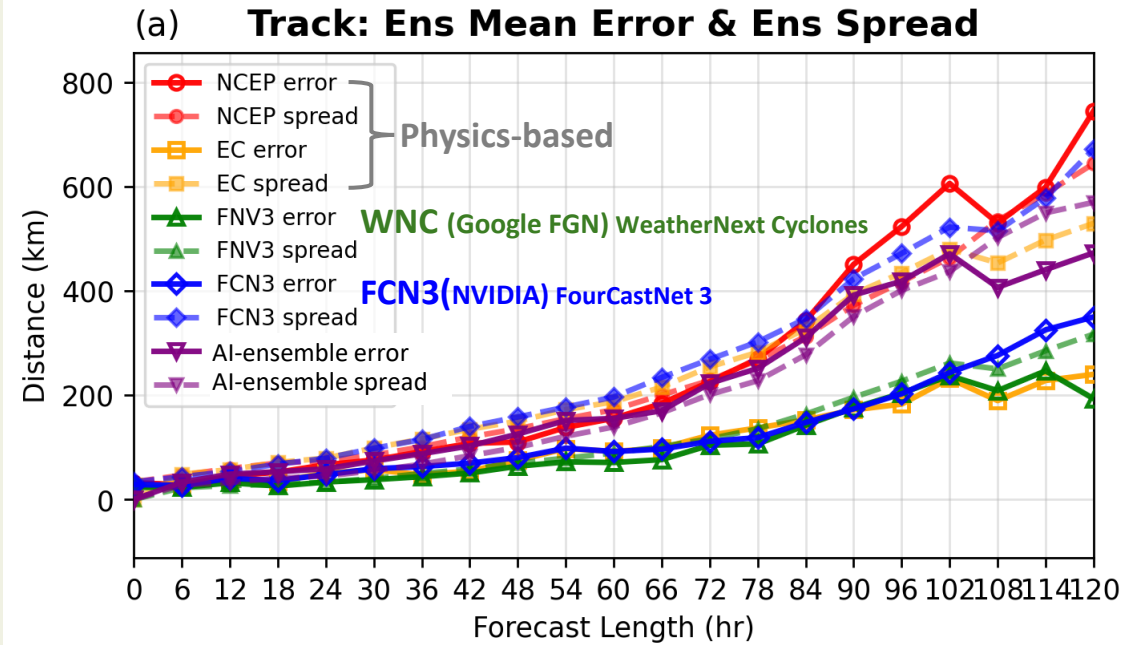
➤ FCN3 forecasts initialized with EC fields generally outperform those initialized with TGFS fields.

FCN3EC **better**/worse than
FCN3TGFS

EMC Verification Scorecard	
Symbol Legend	
▲	32FCNEC is better than 32FCNTGFS at the 99.9% significance level
△	32FCNEC is better than 32FCNTGFS at the 99% significance level
▴	32FCNEC is better than 32FCNTGFS at the 95% significance level
○	No statistically significant difference between 32FCNEC and 32FCNTGFS
▽	32FCNEC is worse than 32FCNTGFS at the 95% significance level
▼	32FCNEC is worse than 32FCNTGFS at the 99% significance level
◊	32FCNEC is worse than 32FCNTGFS at the 99.9% significance level
Not statistically relevant	
Start Date: 2025051600	
End Date: 2025071000	

		Northern Hemisphere							Southern Hemisphere							Tropical						
		day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 1	day 2	day 3	day 4	day 5	day 6	day 7
Anomaly Correlation	Z1000	▲	▲	▲	▲	▲	▲	○	▲	▲	▲	▲	▲	▲	▲	▲	▲	△	△	○	○	○
	Z500	▲	▲	▲	▲	▲	▲	○	▲	▲	▲	△	△	▼	▲	▲	▲	▲	▲	▲	▲	▲
	T850	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	△	▲	▲	▲	▲	▼
	T2M	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
	U10M	▲	△	○	▲	▲	▲	▼	▲	△	▲	△	○	○	○	▼	▼	▼	○	▼	▼	▼
	V10M	△	○	△	○	▲	▲	○	▲	▲	▲	▲	▲	○	▲	▼	▼	○	○	○	○	○
	U850	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	△	○	○	▲	▼	▼	○	○	▲	○	○
	V850	▲	△	△	▲	▲	▲	○	▲	▲	▲	▲	△	▼	▲	▲	▲	▲	△	▲	△	▲
	U250	▲	▲	▲	△	▲	○	○	▲	▲	▲	▲	▲	▲	○	○	○	○	▼	○	○	▲
	V250	▲	▲	▲	▲	▲	△	▼	▲	▲	▲	▲	▲	▼	▲	○	○	○	○	○	▼	▲
RMSE	PPRMSL	▲	▲	▲	▲	▲	▲	○	▲	▲	▲	▲	▲	▲	▲	△	△	△	○	○	○	○
	Z1000	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▲	▲	△	▲	○	○	○	○
	Z500	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	○	△	▲	▲	▲	▲	▲	▲	▲
	T850	▲	▲	▲	▲	▲	▲	▲	▲	△	▲	▲	▲	○	○	▲	○	△	▲	○	▼	▼
	T2M	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
	U10M	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	○	△	△	▼	▼	▼	▼	▼	▼	▼
	V10M	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▼	▼	▼	▼	▼	▼	▼
	U850	▲	▲	▲	▲	△	▲	▼	▲	▲	▲	▲	○	M	▲	▼	▼	▼	△	○	▼	△
	V850	▲	▲	△	○	○	▲	▼	▲	▲	▲	▲	▲	▼	▲	○	▼	▼	▼	▼	○	○
	U250	▲	▲	▲	△	▲	△	○	▲	▲	▲	▲	▲	△	▲	○	▼	▼	▼	▼	○	○
	V250	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	○	▲	▼	▼	▼	▼	▼	▼	○
	PPRMSL	▲	▲	▲	▲	▲	▲	▼	▲	▲	▲	▲	▲	▲	▲	△	△	△	○	○	○	▼

Northwest Pacific Typhoons in July 2025



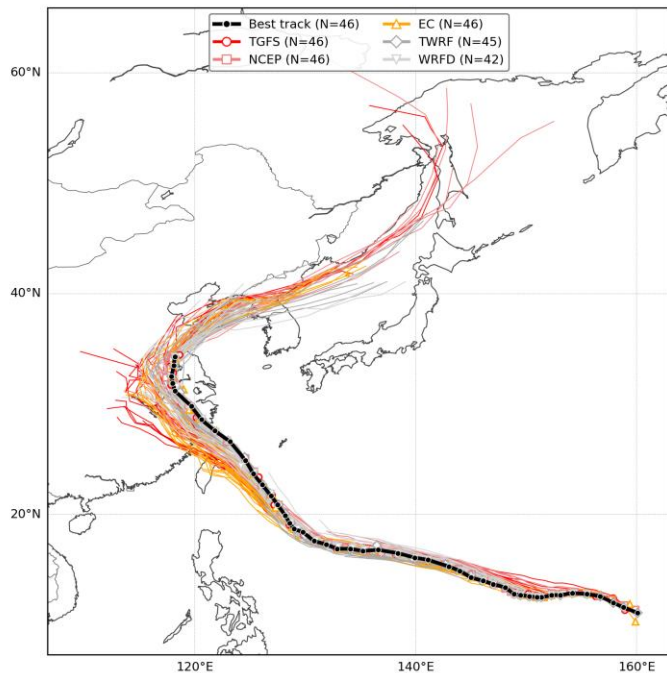
More cases needed for
robust evaluation.

- FCN3 has among the lowest ensemble-mean track errors, slightly behind Google WNC (FGN) but comparable to ECMWF ensemble forecasts.
- FCN3 shows signs of overdispersion; more typhoon cases are needed to confirm this behavior.

Typhoon BAVI (2026)

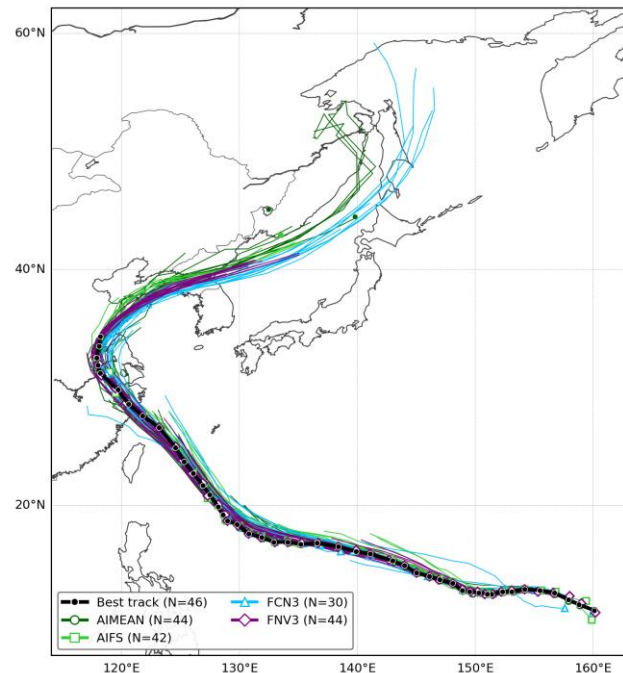
BAVI Deterministic Forecast Tracks

All initialization DTGs: 26070200-26071400 | Forecast: 0-120 h



BAVI Deterministic Forecast Tracks

All initialization DTGs: 26070200-26071400 | Forecast: 0-120 h



- Numerical models:

TWRF > WRFD = NCEP > EC > TGFS

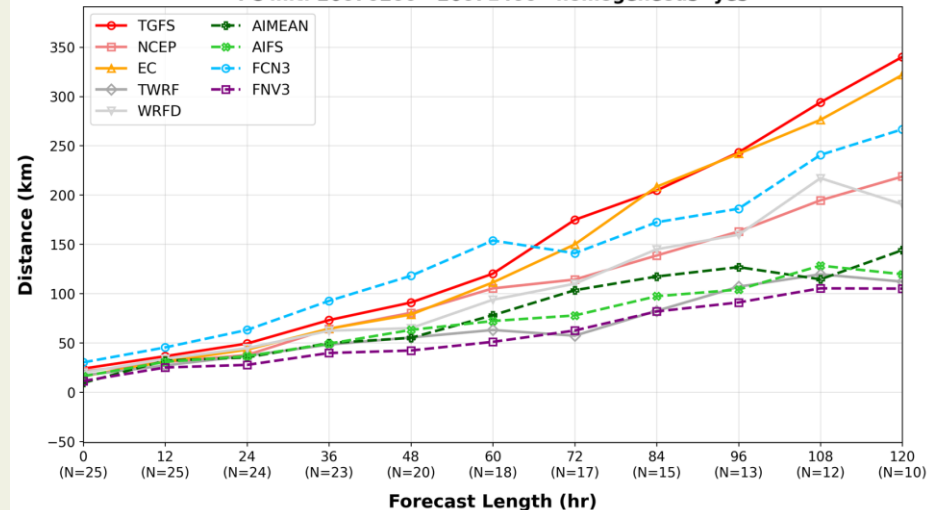
- AI models:

WNC (FGN) > AIFS > AI-ensemble > FCN3

- MLWP models outperform physics-based numerical models in typhoon track forecasting, with WNC showing the best performance.

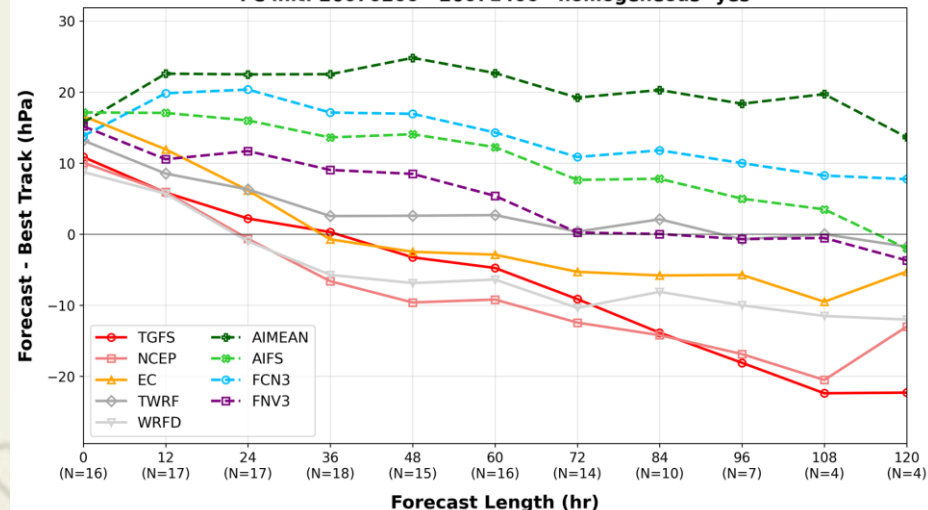
BAVI Deterministic Forecast Track Error

FC init: 26070200 - 26071400 homogeneous=yes



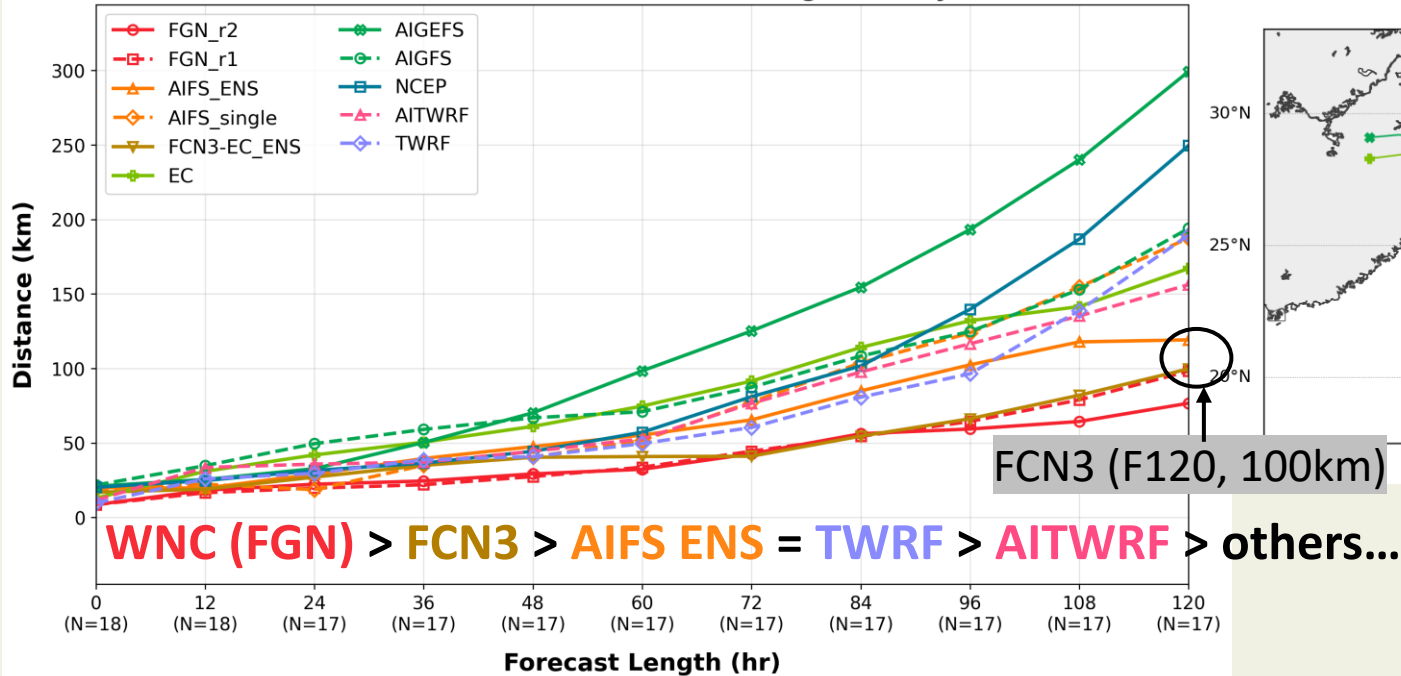
BAVI Deterministic Forecast Minimum SLP Error

FC init: 26070200 - 26071400 homogeneous=yes

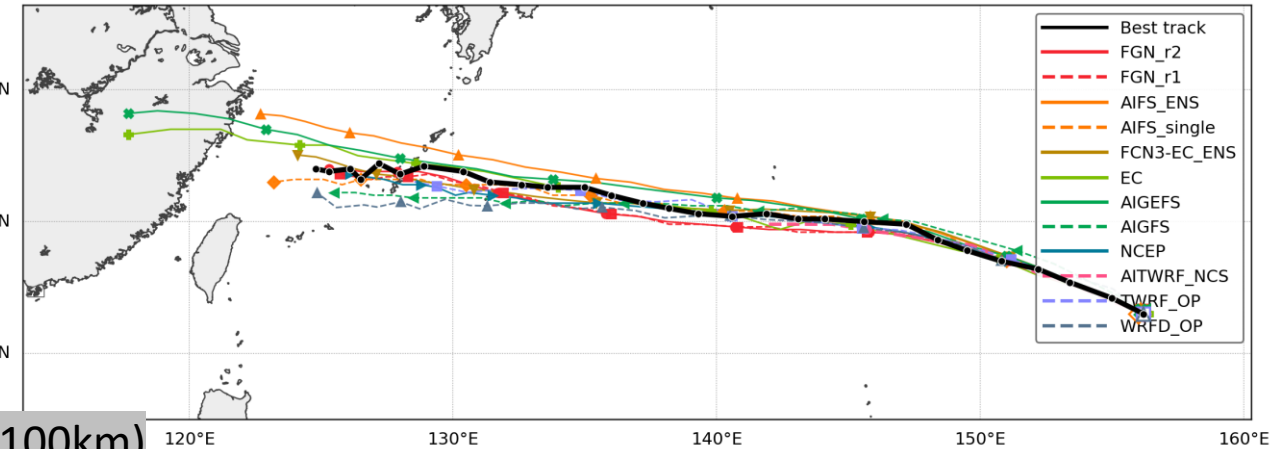


Typhoon DOLPHIN (2026)

DOLPHIN Deterministic Forecast Track Error
FC init: 26072700 - 26081100 homogeneous=yes



DOLPHIN Ensemble-Mean Forecast Tracks
Initialization: 26080112 | Forecast: 0-168 h



- MLWP models show advantages in typhoon track forecasting, while the regional models TWRF and AITWRF also perform well.
- FCN3 performed well in forecasting the track of Typhoon DOLPHIN in 2026, even outperforming the ECMWF physics-based model.

Summary

- Recent MLWP developments highlight a transition toward probabilistic or ensemble-based forecasting for uncertainty representation.
- FCN3 is under evaluation for both weather and subseasonal prediction. Results show good forecast skill and high computational efficiency.
- FCN3 initialized with EC IFS ensemble ICs shows better forecast performance than FCN3 initialized with TGFS ensemble ICs.
- FCN3 shows competitive typhoon track forecast skill, generally comparable to ECMWF ensembles in July 2025.
- Related talk: FCN3's performance on subseasonal timescales. (A3-1郭珮萱)

A world map with a stylized, circuit-like background. The map is overlaid with a network of glowing yellow lines and dots, suggesting global connectivity. The background features a light gray circuit pattern and several large, semi-transparent brown circles in the lower right corner.

**Thanks for
Your Attention!**



Central Weather Administration 2026