Verification of Tropical Cyclone Operational Forecast in 2025

BAI Lina, YANG Mengqi, CHEN Guomin

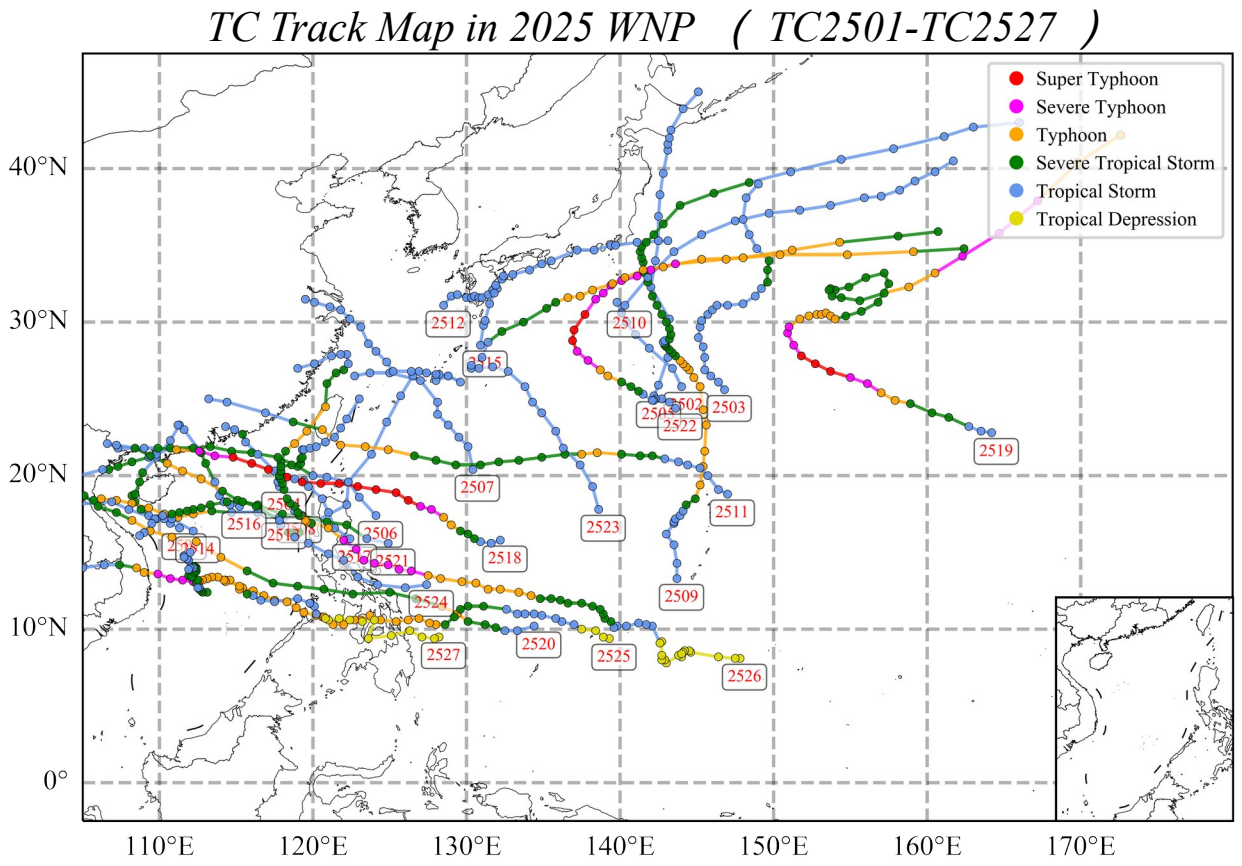
(Shanghai Typhoon Institute/CMA)

10-13 March 2026,
Jeju Island, Republic of Korea

Outline

1. Data
2. Performance of TC track forecasts
3. Performance of TC intensity forecasts
4. Conclusion

Data TC reference data



TC ID	English Name
2501	WUTIP
2502	SEPAT
2503	MUN
2504	DANAS
2505	NARI
2506	WIPHA
2507	FRANCISCO
2508	CO-MAY
2509	KROSA
2510	BAILU
2511	PODUL
2512	LINGLING
2513	KAJIKI
2514	NONGFA
2515	PEIPAH
2516	TAPAH
2517	MITAG
2518	RAGASA
2519	NEOGURI
2520	BUALOI
2521	MATMO
2522	HALONG
2523	NAKRI
2524	FENGSHEN
2525	KALMAEGI
2526	FUNG-WONG
2527	KOTO

- TC reference data ：
  - *Different TC best track datasets or real-time operational TC position and intensity datasets were used as reference data to verify official guidances.*
  - *TC best track dataset from the RSMC-Tokyo Center was used as the reference dataset to verify objective forecasts.*

Data

TC position and intensity forecast data

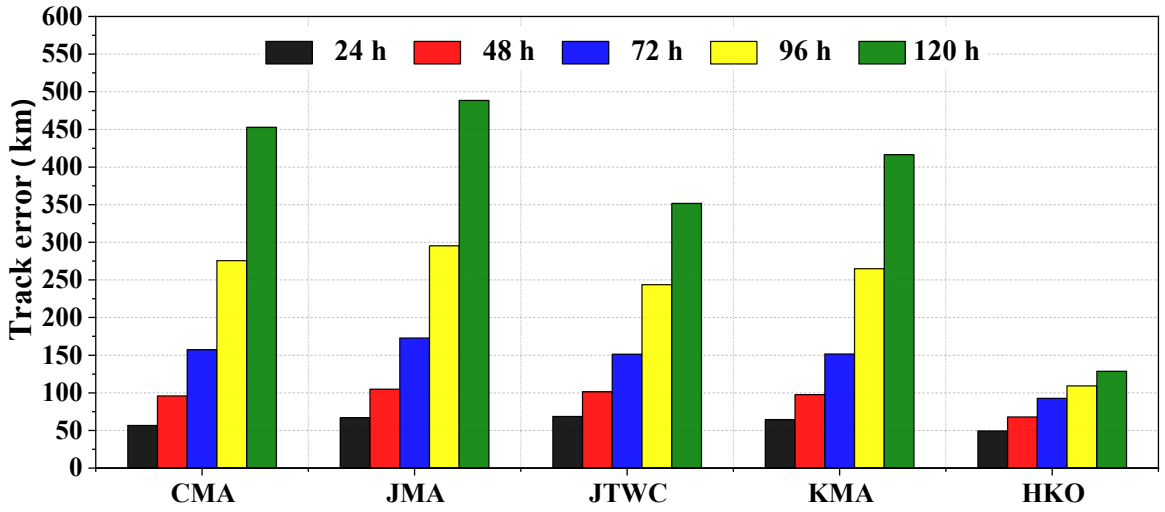
Operational official guidances, NWP models, AIWP models

Category	Abbreviation		Full name or description	Source
Deterministic	***Official guidance***	**CMA**	*China Meteorological Administration*	CMA
		JMA	*Japan Meteorological Agency*	JMA
		JTWC	*Joint Typhoon Warning Center*	JTWC
		KMA	*Korea Meteorological Administration*	KMA
		HKO	*Hong Kong Observatory*	HKO
	Global NWP Model	**CMA-GFS**	*CMA Global Forecast System*	CMA
		ECMWF-IFS	*ECMWF Integrated Forecasting System*	ECMWF
		JMA-GSM	*JMA Global Spectral Model*	JMA
		NCEP-GFS	*NCEP Global Forecast System*	NCEP
		UKMO-MetUM	*Met Office Unified Model*	UKMO
	Regional NWP Model	**CMA-TYM**	*CMA Regional Typhoon Forecasting Model*	CMA
		CMA-TRAMS	*CMA Tropical Regional Atmospheric Model System of the South China Sea*	ITMM/CMA
		SWARMS-TC	*Shanghai Weather and Risk Model System–Typhoon Model*	STI/CMA
	AIWP Model	**FengQing**	*Artificial Intelligence Global Short and Medium Range Forecasting System*	CMA, THU
		FengWu	*Artificial Intelligence Global Weather Forecasting Model*	Shanghai AI Lab
		FuXi	*Machine Learning-Based Global Weather Forecasting Model*	Fudan University
		Pangu	*Artificial Intelligence Global Weather Forecasting Model*	Huawei Cloud
		AIFS	*Artificial Intelligence Forecasting System*	ECMWF
		GraphCast	*Machine Learning-Based Global Weather Forecasting Model*	Google DeepMind
Ensemble	***NWP Model***	**ECMWF-EPS**	*ECMWF Ensemble Prediction System*	ECMWF
		JMA-GEPS	*JMA Global Ensemble Prediction System*	JMA
		NCEP-GEFS	*NCEP Global Ensemble Forecast System*	NCEP
		UKMO-EPS	*UKMO Ensemble Prediction System*	UKMO
		SWARMS-EN	*Shanghai Weather and Risk Model System–Ensemble Typhoon Model*	STI/CMA

Performance of TC track forecasts

Official guidances

Track Errors (2025)



Official guidances **mean track error** (2024 vs. 2025)

	2024	2025	Decrease rate
24h	≈70 km	≈60 km	**14%**
48h	≈105 km	≈95 km	**10%**
72h	≈160 km	≈150 km	**6%**
96h	≈260 km	≈250 km	**4%**
120h	≈420 km	≈405 km	**4%**

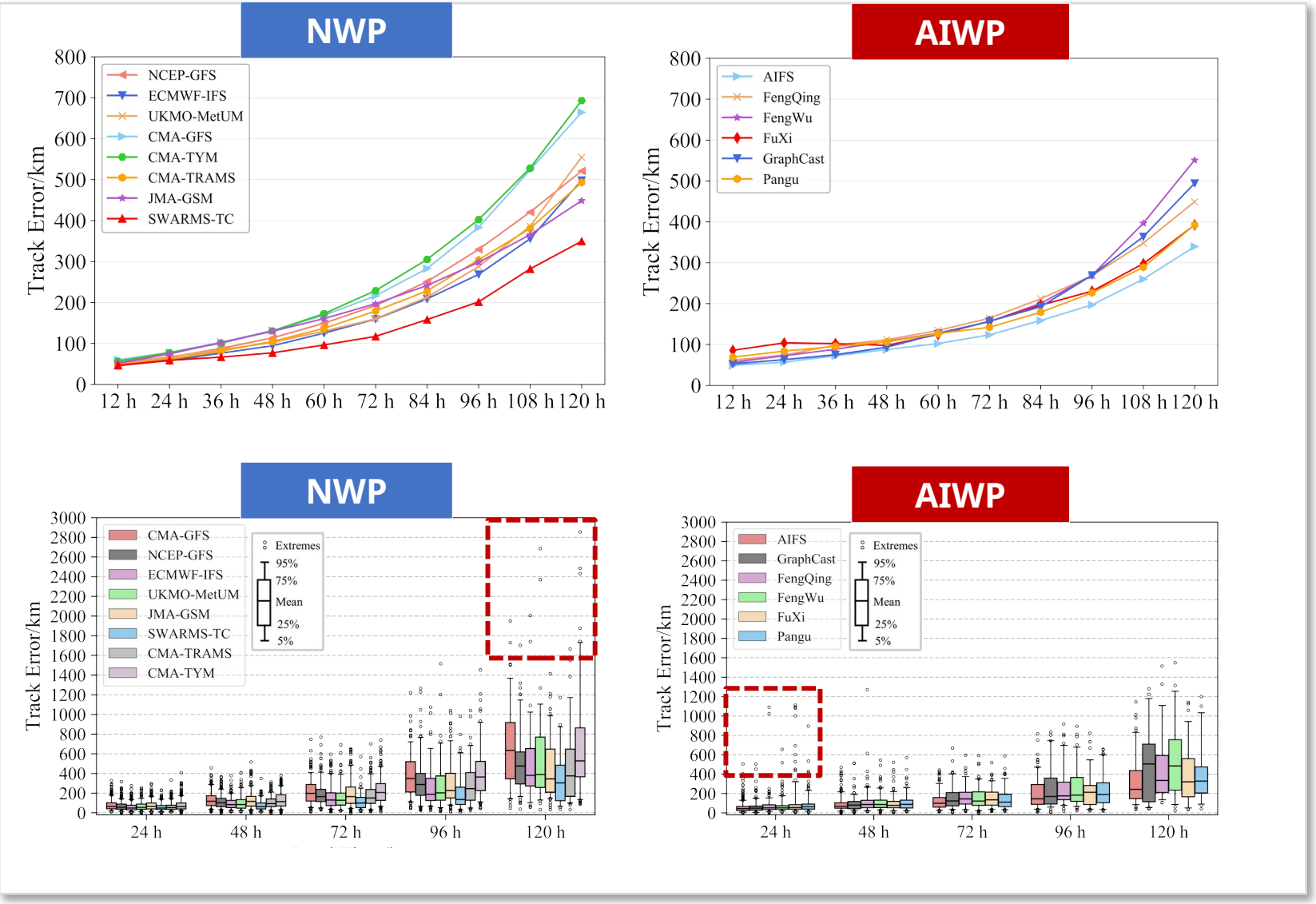
Track errors ：

- The overall mean track errors decreased in 2025 compared to 2024.
- Notably, the overall mean track error at 24 h lead time was reduced by 14%.

Performance of TC track forecasts

Deterministic NWP and AIWP

Track Errors (2025)



NWP Mean Track Errors:

24 h: 60~80 km , 48 h: 80~130 km ,

72 h: 120~230 km , 96 h: 200~400 km , 120 h: 350~700 km

For NWP models, the difference in the mean track errors increased as lead times increased.

AIWP Mean Track Errors:

24 h: 60~100 km , 48 h: 90~110 km ,

72 h: 120~165 km , 96 h: 200~270 km , 120 h: 340~550 km

For AIWP models, the smallest difference in mean track errors occurred at 48 h, while the difference gradually increased beyond 48 h.

Large errors were associated with the westward-bias of Super Typhoon Halong (TC2522).

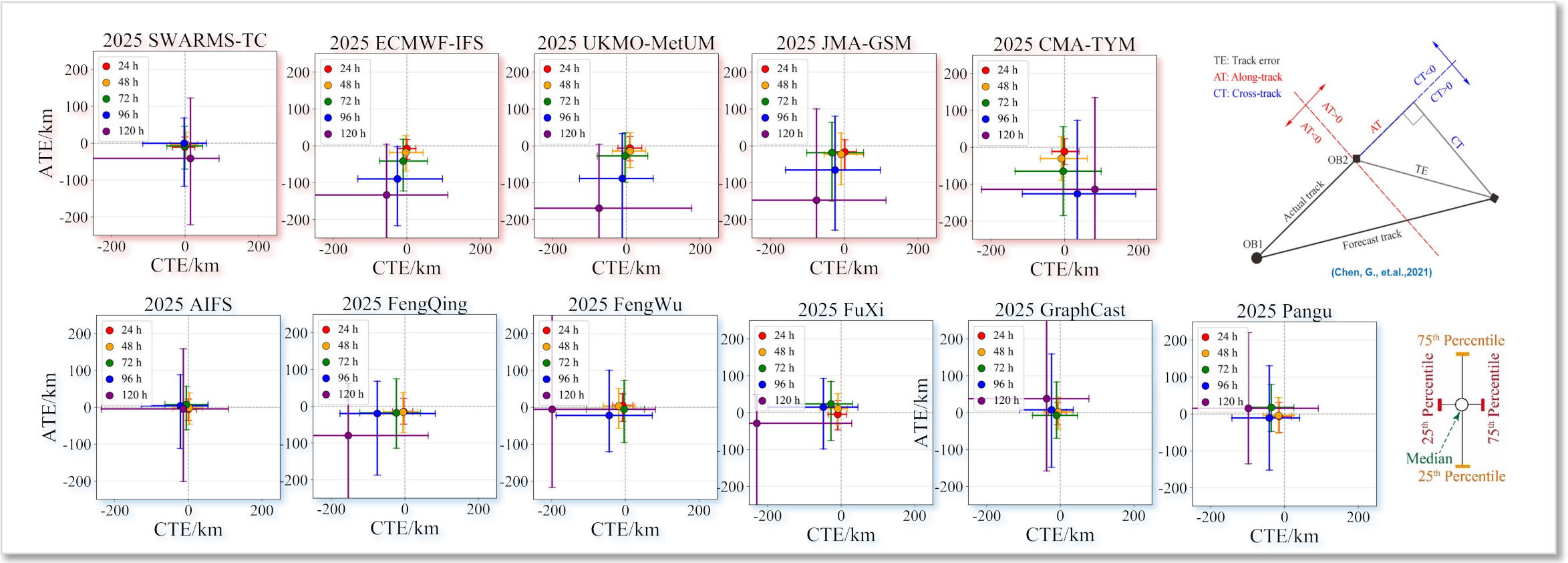
Performance of TC track forecasts

Deterministic NWP and AIWP

ATE & CTE (2025)

NWP

AIWP

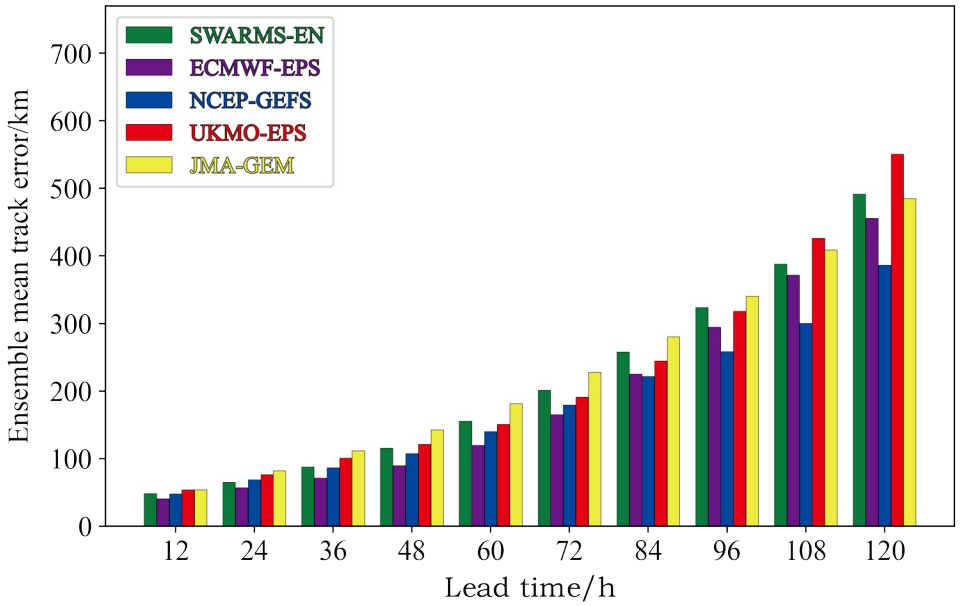


ATCT Bias ：

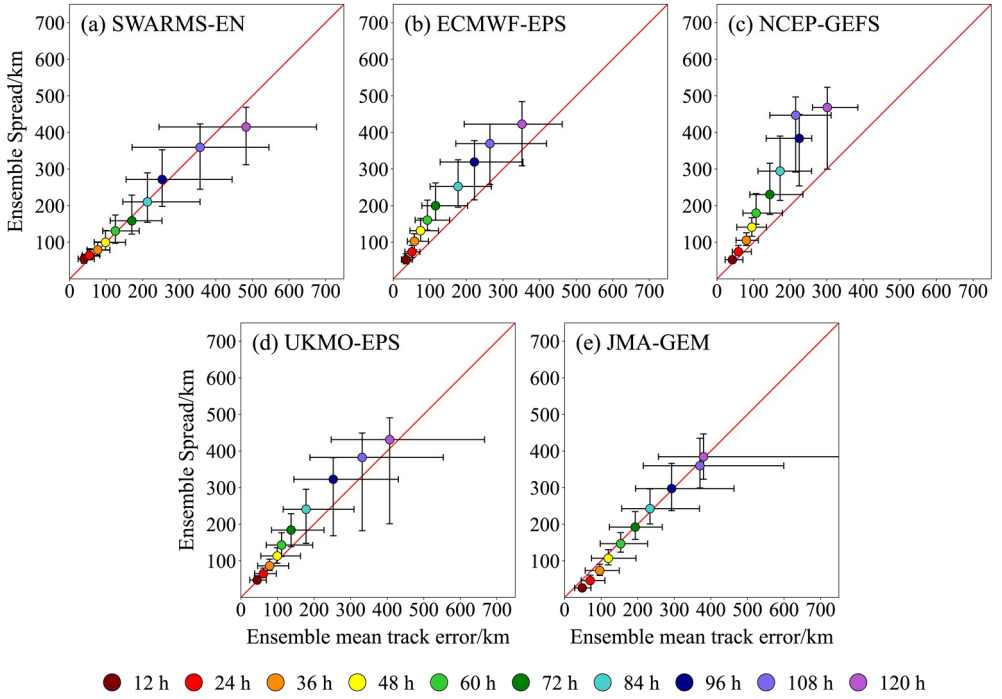
- Traditional NWP models exhibit pronounced slow bias, while ML-physics hybrid models (SWARMS-TC) exhibit small bias.
- AIWP models show varying systematic biases, with AIFS exhibiting the smallest bias.

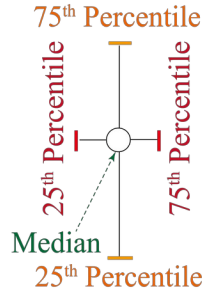
Performance of TC track forecasts

Ensemble NWP models



Joint plot of ensemble mean position error and ensemble spread





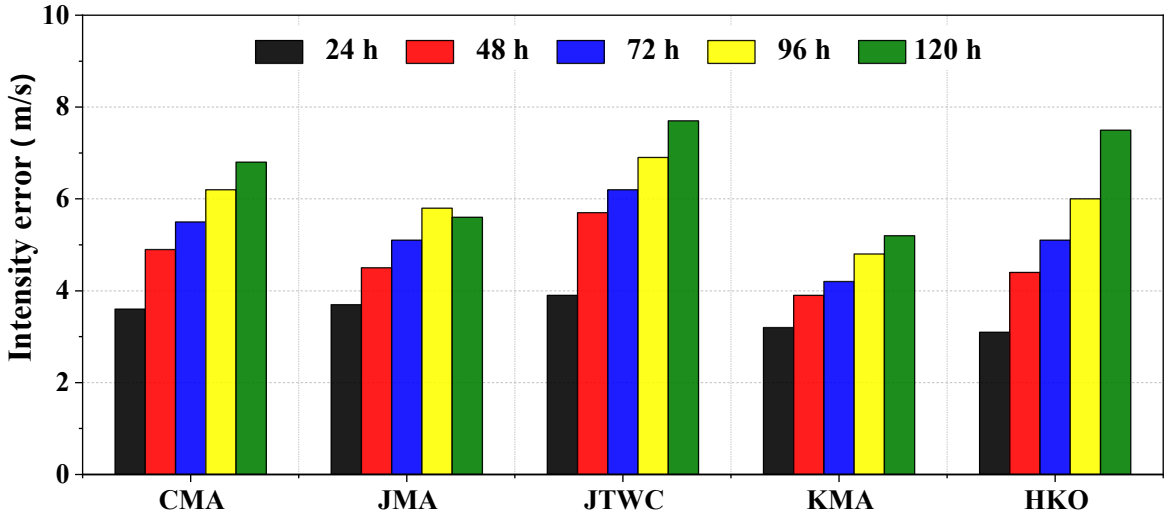
In 2025 ：

- JMA-GEM and SWARMS-EN appropriately expressed track forecast uncertainty.
- While ECMWF-EPS and NCEP-GEFS showed smaller mean errors, forecast uncertainties grew as the lead time increased.

Performance of TC intensity forecasts

Official guidances

Intensity Errors (2025)



Official guidances

Mean Absolute Errors (2024 vs. 2025)

	2024	2025	Decrease rate
24h	≈4 m/s	≈3.5 m/s	**13%**
48h	≈5.5 m/s	≈4.7 m/s	**15%**
72h	≈6 m/s	≈5.2 m/s	**13%**
96h	≈7.5 m/s	≈5.9 m/s	**21%**
120h	≈10 m/s	≈6.4 m/s	**36%**

In 2025 ：

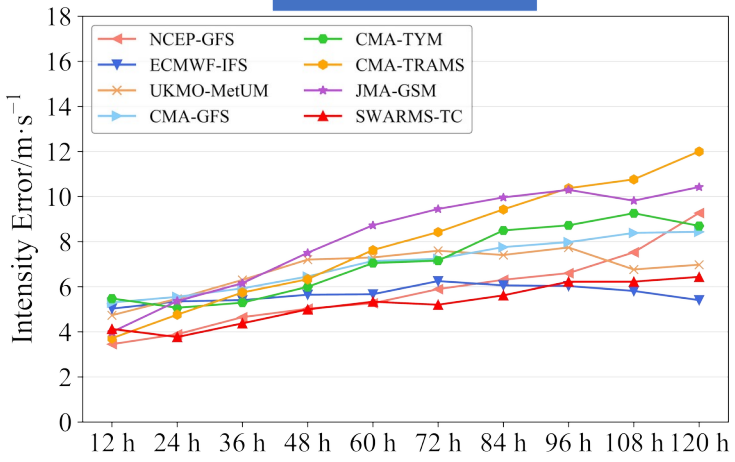
- Intensity errors of official guidances gradually increased as lead times increased.
- Mean absolute intensity errors gradually increased from 3 ~ 4 m/s at 24h to 5 ~ 8 m/s at 120h.
- The overall mean errors decreased in 2025 compared to 2024.

Performance of TC intensity forecasts

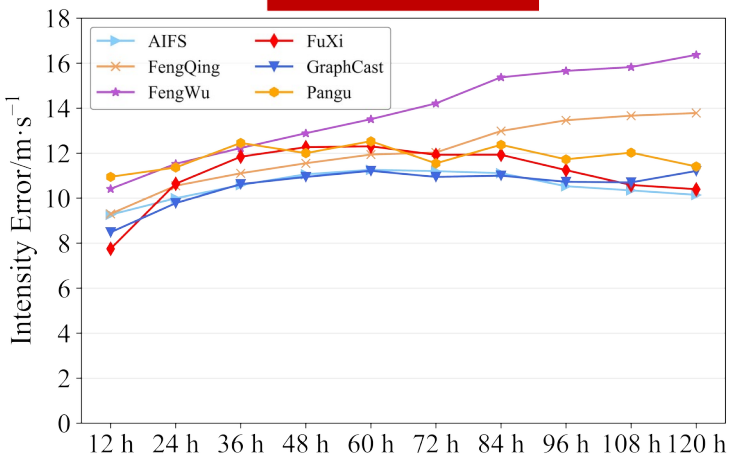
Deterministic NWP and AIWP

Intensity Errors (2025)

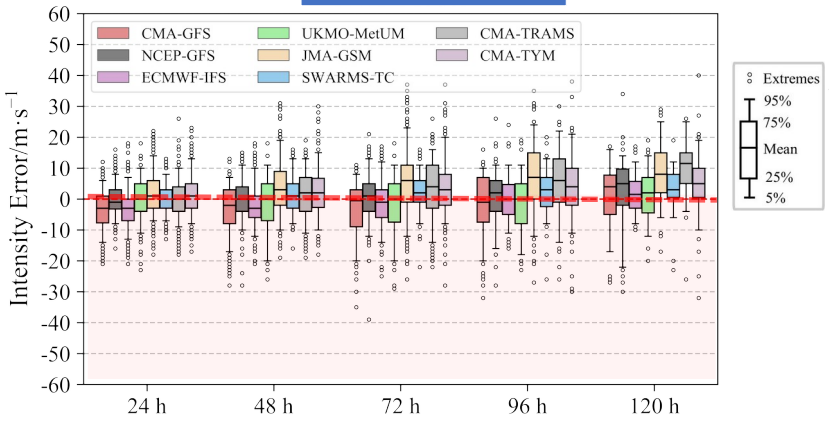
NWP



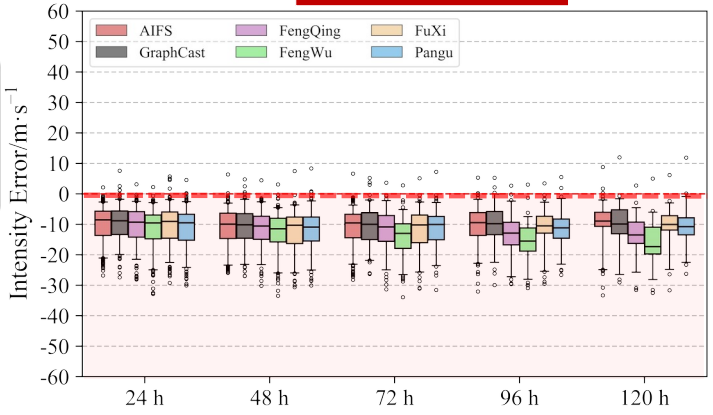
AIWP



NWP



AIWP



NWP Mean Absolute Errors (MAEs) :

24 h: 4~6 m/s ， 120 h: 5~12 m/s

- For NWP models, the difference in MAEs increased with lead times, showing tendency to overestimate intensity at longer lead times.

AIWP MAE:

24 h: 10~12 m/s ， 120 h: 10~16 m/s.

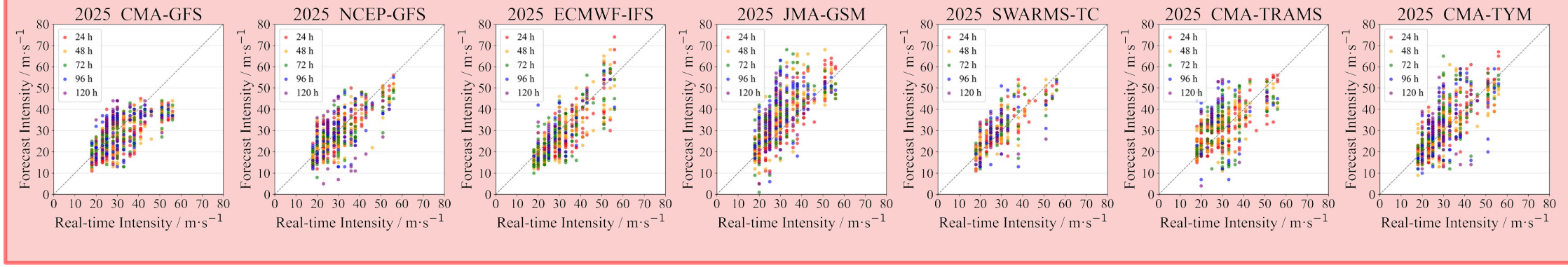
- The MAEs for AIWP were larger than NWP.
- AIWP models underestimated intensity, with only a few cases overestimating intensity.
- The degree of underestimation showed improvement compared to 2024.

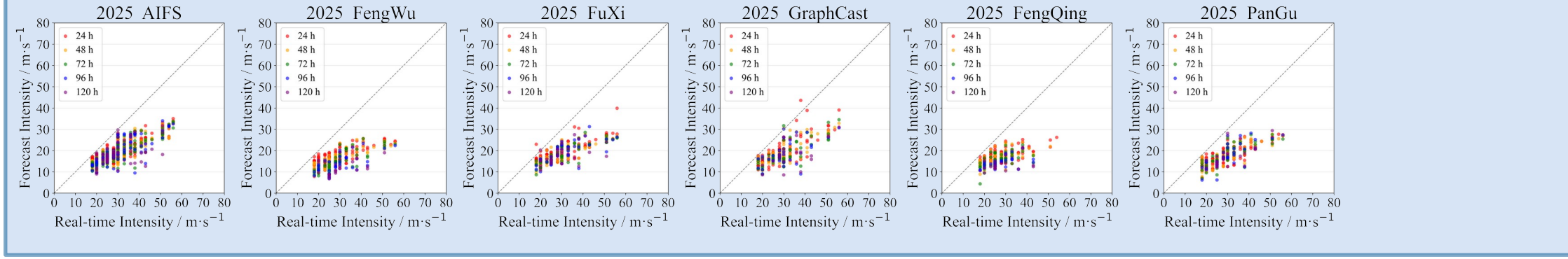
24 h	Forecast method	Rate of Intensity Consistent Trend (RCT)
NWP	NWP	60~70 %
AIWP	AIFS Other AIWP models	60% 40~50%

Performance of TC intensity forecasts

Deterministic NWP and AIWP

Forecast Intensity vs Real-time Intensity (2025)





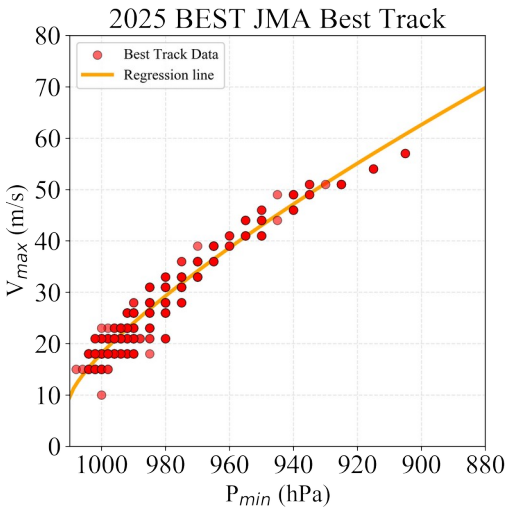
NWP ： NWP underestimated or overestimated TC intensity at different degree.

AIWP: The stronger the intensity, the more the AIWP tended to underestimate it.

Performance of TC intensity forecasts

Deterministic NWP and AIWP

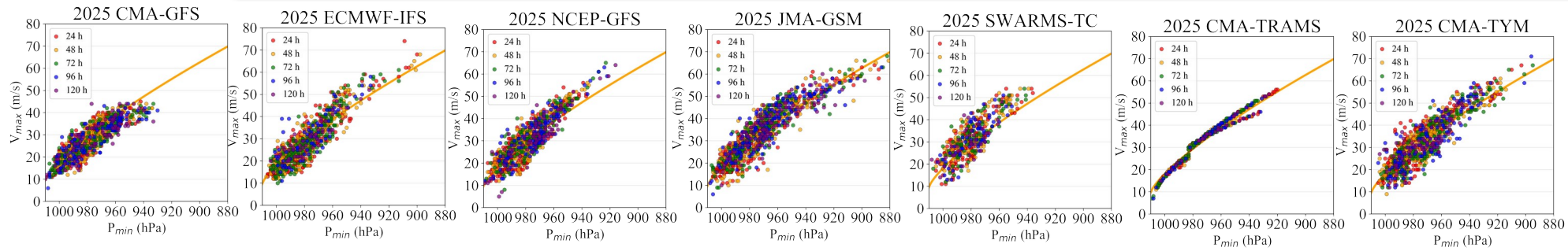
Wind-Pressure Relationship: Forecast Wind Speed vs. Forecast Pressure



Wind-Pressure Relationship

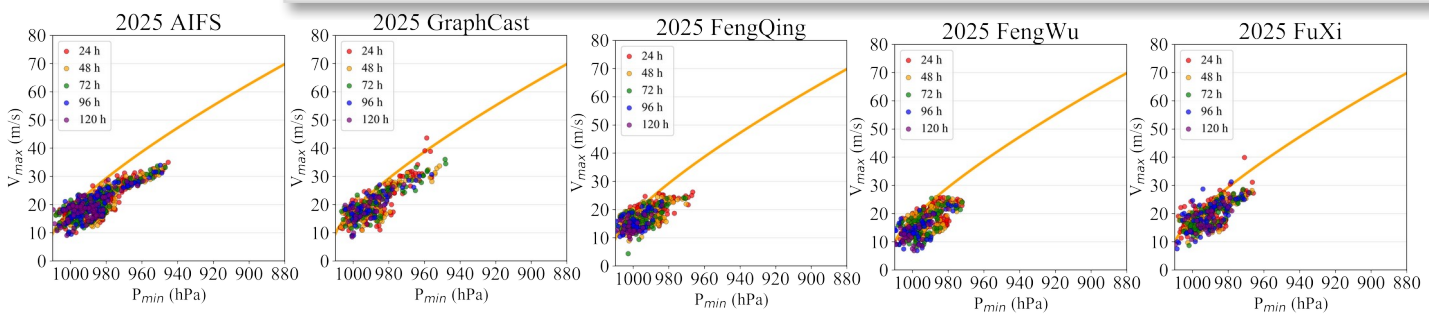
➢ NWP

Wind-pressure relationship forecast by NWP was more consistent with observations.



➢ AIWP

AIWP models exhibit systematic biases in the wind-pressure relationship.



Mechanism Differences:

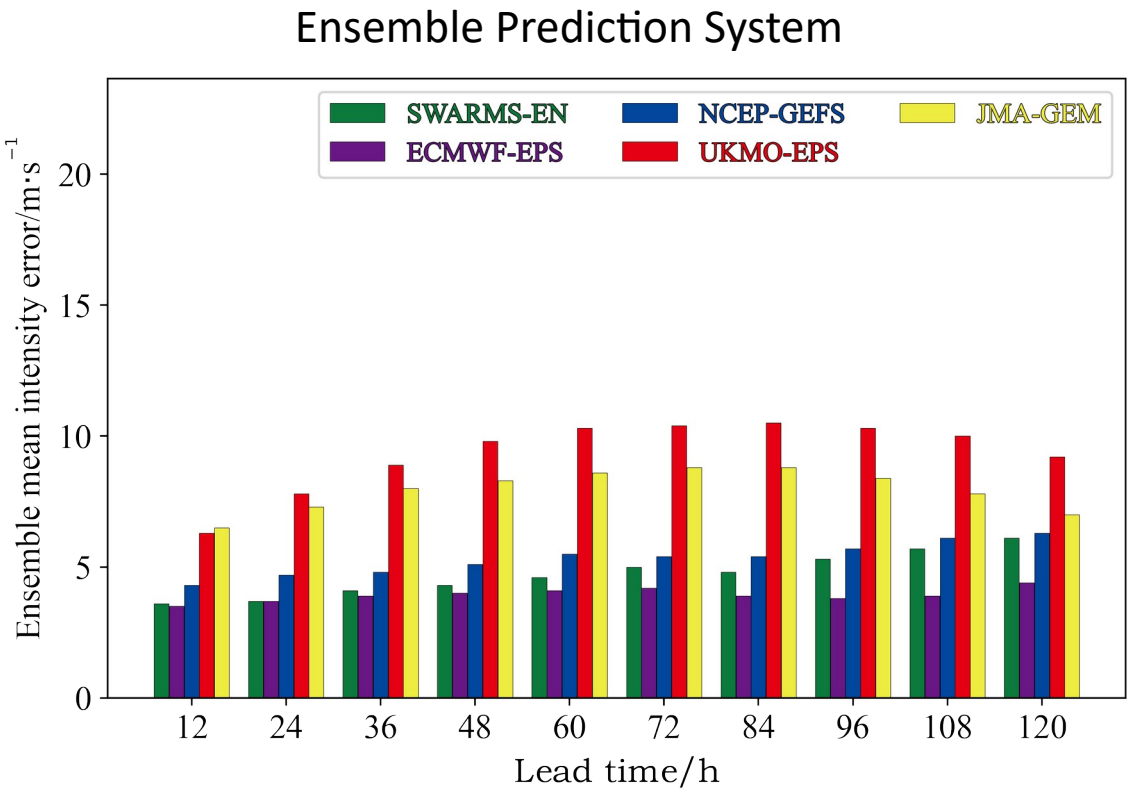
- NWP : Atmospheric motion equations, maintaining physical constraints.
- AIWP: Statistical fitting, lacking physical constraints.

Limitations of Training Samples:

- Data scarcity (few samples of intense TCs);
- AI models struggle to learn the nonlinear characteristics of the wind-pressure relationship, leading to underestimation of intensity.

Performance of TC intensity forecasts

Ensemble NWP models



In 2025 ：

- Among the Ensemble Prediction Systems (EPSs), the ECMWF-EPS exhibited the smallest ensemble mean errors, with mean errors of approximately 4 m/s, followed by the NCEP-GFS and SWARMS-EN.

Released TC Forecast Verification Dataset

Tropical Cyclone Forecast Verification Dataset

Released by the team working on Typhoon Committee's Perennial Operating Plan (POP) 'Verification of Tropical Cyclone Operational Forecast'

Data Content

The dataset covers TC track and intensity forecast error over the western North Pacific and the South China Sea.

95th Error : Forecast errors ≥ 95th percentile error in its group (grouped by year, forecast method, and forecast lead time).

Data Access : CMA TROPICAL CYCLONE DATA CENTER for the western North Pacific Basin

https://tcdata.typhoon.org.cn/en/ybjdpd.html

01 Forecast Method

① **Global/Regional NWP Model :**
SWARMS-TC, CMA-GFS, ECMWF-IFS, etc.

② **AIWP Model :**
Fengqing, FengWu, FuXi, Pangu, etc.

02 Data Column

① **TC ID**

② **Initial forecast time (UTC)**

③ **Forecast lead time**

④ **Forecast data :**
TC longitude, latitude, intensity

⑤ **Forecast error :**
position error, intensity error

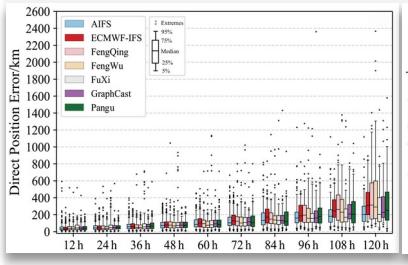
03 Release & Plan

① **First release :**
1 June 2025 (2024 dataset)

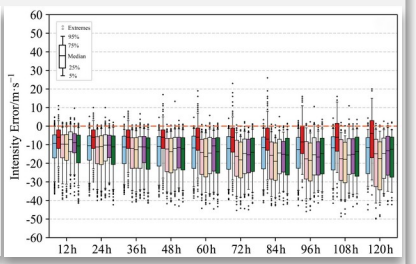
② **Update frequency :**
updated annually, typically in the first half of the following year

③ **Future plan :**
extension to wind and precipitation forecasts
extension to landfall forecasts
extension to global TC regions

04



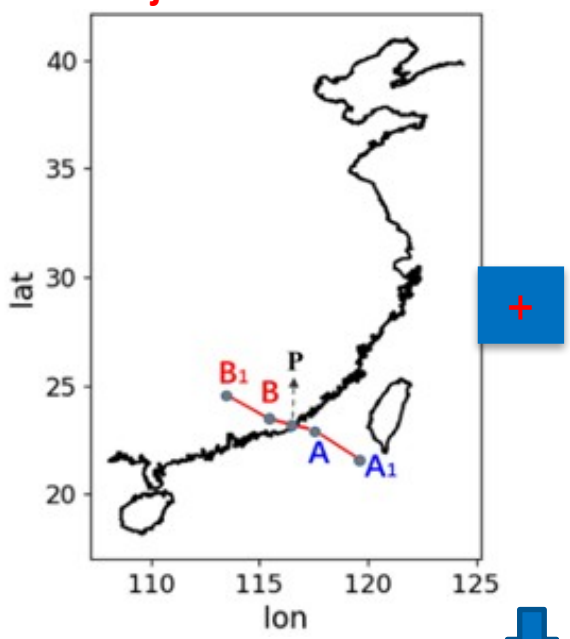
2024 TC position error



2024 TC intensity error

Future plan- verification of TC landfall forecasts

Objective Method



Definition of Typhoon Landfall Locations

Land country	Remarks
China	including Taiwan, Hong Kong, Hainan, Zhoushan, Chongming, Pingtan, Dongshan, Nanao, Donghai, Nansan, Hailing, Zhujiajian, Jintang, Daishan, Gaotang, Nantian, Xilian, Daishan, and Dongtou islands. (19) CMA's 'Regulations on Typhoon Operations and Services'
Japan	Excluding islands other than the mainland of Hokkaido, Honshu, Shikoku, or Kyushu. (4) JMA's website: https://www.data.jma.go.jp/typhoon/statistics/landing/landing.html
Philippines	Including all islands in Philippine. Annual Report on Philippine Tropcical Cyclones 2020
Vietnam	
Thailand	
Republic of Korea	Including Jeju, Ulleungdo, Geojedo, Wando, Jindo, Ganghwado, Namhae, and Aphaedo islands. **(Provided by the National Typhoon Center / KMA)** (8)
North Korea	
Malaysia	
United States (Guam)	

Landfalling datasets for the Asia-Pacific region based on objective method

	A	B	C	D	E	F	G	H	I	J
1	TCNUM	TCID	YEAR	NAME	LANDFALL_TIME	LAT(degree_north)	LON(degree_east)	WIND(m/s)	PRESSURE(hPa)	INT_CAT
2	3	9103	1991	vanessa	1991-04-28 13:32:19	18.6615	110.262	15	999	1
3	7	9106	1991	zeke	1991-07-13 05:36:30	18.6601	110.259	42	963	5
4	8	9107	1991	amy	1991-07-19 16:28:07	23.1645	116.642	39	955	4
5	9	9108	1991	brendan	1991-07-24 07:36:11	21.9669	113.259	30	980	3
6	13	9111	1991	fred	1991-08-16 07:33:28	20.4	110.481	45	960	5
7	18	9116	1991	joel	1991-09-06 22:59:28	22.7993	115.101	28	982	3
8	22	9119	1991	nat	1991-09-23 06:33:04	22.2	120.893	46	948	5
9	4	9204	1992	chuck	1992-06-28 07:09:44	18.4162	109.912	35	961	4
10	6	9205	1992	eli	1992-07-13 07:59:33	19.1988	110.602	35	970	4
11	7	9206	1992	faye	1992-07-18 05:42:14	22.1087	113.462	20	992	2
12	8	9207	1992	gary	1992-07-23 00:25:44	20.9165	110.636	31	979	3
13	14	9212	1992	mark	1992-08-19 07:11:47	23.5598	117.094	20	993	2
14	16	9215	1992	omar	1992-09-04 20:38:03	23.6317	121.531	30	976	3
15	17	9216	1992	polly	1992-08-30 12:10:34	23.9784	121.622	32	975	3
16	21	0	1992	nameless	1992-09-19 15:43:08	19.0713	110.57	13	1002	1

A comprehensive set of typhoon landfall records integrating the data officially released by the meteorological departments of each country

Verification of TC landfall forecasts

Conclusion

Track forecast ：

- *Traditional NWP models exhibit pronounced slow bias, ML-physics hybrid models (SWARMS-TC) exhibiting small bias.*
- *For AIWP models, the smallest difference in mean track errors occurred at 48 h lead time, while the difference gradually increased beyond 48 h. AIWP models show varying systematic biases, with AIFS exhibiting the smallest bias.*

Intensity forecast ：

- *For NWP models, NCEP-GFS (at lead times ≤ 96 h) and SWARMS-TC maintained comparatively lower MAEs with a relatively small increase in MAEs as the lead time increased.*
- *AIWP models consistently showed inferior performance compared to NWP models, with significant underestimation across all lead times and low consistency in intensity trends. However, the degree of underestimation showed improvement compared to 2024.*

Thanks

bailn@typhoon.org.cn