Typhoon Committee Working Group on Meteorology

Annual Operating Plan 13: Promoting Technical Exchange of AI Applications in Tropical Cyclone Analysis and Forecasting

Activity Report 2025

prepared by the Expert Team on AI Applications in Tropical Cyclone Analysis and Forecasting (ET-AITC)

1. Introduction

The field of weather forecasting is undergoing a potentially profound transformation driven by advances in artificial intelligence and data-driven modeling. Traditional numerical weather prediction (NWP) systems rely on physics-based simulations that demand vast computational resources and often face limits in the detailed physical representations of many atmospheric processes. In contrast, AI-based models such as FourCastNet, Pangu-Weather, GraphCast, and AIFS have demonstrated the ability to learn from decades of atmospheric data and generate forecasts that rival or even surpass conventional methods, while being orders of magnitude faster. Published in leading peer-reviewed journals, these models leverage deep learning architectures to capture complex atmospheric dynamics with remarkable efficiency and accuracy.

One of the most compelling applications of this AI technology is in tropical cyclone forecasting, where timely and accurate predictions can save lives and guide disaster response. AI-driven models have shown promise in predicting more accurate TC tracks than traditional NWP systems, particularly over longer lead times. By cutting-edge machine learning techniques, these systems are beginning to reshape the TC forecasting landscape, offering the potential for better TC track guidance.

1.1. Kick-start workshop

The Hong Kong Observatory (HKO) organised a workshop on "Promoting Technical Exchange of Artificial Intelligence (AI) Applications in Tropical Cyclone Analysis and Forecasting" under the United Nations Economic and Social Commission for Asia and the Pacific (ESCAP) / World Meteorological Organization (WMO) Typhoon Committee on 21 to 22 May 2024.

29 experts from 11 Typhoon Committee Members, namely China; Hong Kong, China; Japan; Lao People's Democratic Republic; Macao, China; Malaysia; the Philippines; Republic of Korea; Thailand; Socialist Republic of Viet Nam and the United States of America attended the workshop. Participants engaged in in-depth discussions and technical sharing on the developments and applications of AI in tropical cyclone analysis and forecasting. Their presentations covered objective analysis techniques of tropical cyclone, monitoring and forecasting extreme weather events, applications of data-driven models as well as forecast verification. The workshop also invited leading AI development teams from World Meteorological Centre, academia and information technology company to share the latest advancements in AI applications in weather forecasting.

The workshop discussed user requirements for AI technologies and products within the Typhoon Committee region, as well as potential future collaboration projects. In an effort to strengthen the capabilities of tropical cyclone monitoring, forecasting and warning, the Workshop recommended that (1) a platform could be developed to facilitate Typhoon Committee Members to share, use and develop AI applications in TC analysis and prediction; (2) good practices or points to note could be shared to shorten the learning curve, including to set up an expert team under WGM; (3) more training and capacity building activities such as future workshops in series.

1.2. Launch of WGM AOP13 and establishment of the expert team

Following the recommendations from the kick-start Workshop and subsequent discussions in the 7th annual meeting of WGM, HKO proposed an AOP, initially being a Research-to-Operation project in 2025, to continue the initiative among interested TC Members. Upon endorsement at the 57th Session of the Typhoon Committee (TC57), the Expert Team on AI Applications in Tropical Cyclone Analysis and Forecasting (ET-AITC) under TC WGM was established to collaboratively lead the activities of this AOP in 2025 and beyond.

The Expert Team will lead the WGM's AOP13, continuously assess and review AI-based methods in tropical cyclone analysis and forecasting, and coordinate the data exchange and verification of AI-generated TC forecast products among TC Members. It will compile and submit an annual report on AI model development, applications, and performance, with a focus on cyclone track forecasting, while supporting workshops, coordinating with capacity-building activities, and periodically review the AOP scope to reflect latest progress and emerging needs.

Members of the Expert Team are listed in Appendix 1.

2. Member reports: The development, utilization, and future plans of AI technology for typhoon analysis and forecasting

2.1 China

In recent years, with the rapid advancement of artificial intelligence technology, the National Meteorological Center (NMC) of the China Meteorological Administration (CMA) has deeply integrated AI into typhoon operations and established an objective AI-based typhoon monitoring and forecasting product system. This system covers the entire process from typhoon vortex identification and intensity estimation to track and intensity prediction, improving the accuracy and timeliness of typhoon forecasting and providing technical support for disaster prevention and mitigation. The main progress of technologies are as follows:

2.1.1 Typhoon Vortex Identification Technology

The NMC has developed a multi-scale iterative satellite image target detecting system based on Single Shot Multi-Box Detector (SSD) to address the challenge of identifying typhoon vortices. This technology employs a two-step strategy of coarse and fine typhoon vortex positioning, enabling efficient and accurate typhoon vortex extraction from infrared cloud imagery. Evaluation results show that the identification rate for targets of severe tropical storm intensity and above (wind speed of Beaufort Scale 10 or higher) reaches nearly 95%, while the identification rate for tropical depressions and tropical storms ranges from 40% to 90%.

2.1.2 Typhoon Intensity Estimation Model

After accurately identifying the typhoon vortex, the results are applied to an AI-based typhoon intensity estimation model. This model is a fusion model based on dual tasks of image classification and regression. It uses a convolutional neural network (CNN) deep learning model to extract features related to typhoon intensity from satellite cloud imagery. Based on these features, a classification model and a similarity-based retrieval model are constructed, and the final intensity and confidence level are derived by integrating the results of both models. Test results from 2020 show that the model's mean absolute error (MAE) and root mean square error (RMSE) for typhoon intensity estimation are 3.9 m/s and 5.4 m/s respectively, demonstrating strong reference value for operational typhoon intensity estimation.

2.1.3 Joint Estimation of Typhoon Scale and Intensity—DeepTCNet

The DeepTCNet model, developed by the team of Professor Zhemin Tan from Nanjing University, has been deployed and is operational at the NMC. This model uses a physics-enhanced deep learning approach to estimate the intensity and scale of tropical cyclones from satellite infrared imagery. In addition to satellite cloud imagery, the model innovatively incorporates physical information such as typhoon "fullness," as well as typhoon location and movement, as joint input data. Preliminary evaluations indicate that its performance in estimating typhoon intensity is superior to the Advanced Dvorak Technique (ADT), and its scale estimation outperforms result diagnosed from Multiplatform Tropical Cyclone Surface Wind Analysis (MTCSWA).

2.1.4 Typhoon Rapid Intensification Discrimination Model

To address the common challenge of forecasting rapid intensification (RI) of typhoons, the NMC has developed a typhoon RI discrimination model that integrates spatiotemporal sequence features. This model can predict whether a typhoon will undergo RI within the next 12 hours using satellite cloud imagery. The comprehensive accuracy (ACC) of the model for RI prediction exceeds 75%, providing forecasters with critical decision-making support.

2.1.5 TYTEC-AI

The NMC has reconstructed the TYTEC technology using deep learning, resulting in the TYTEC-AI technology. This technology employs an adaptive ensemble correction technique, which uses real-time evaluation results of ensemble numerical weather model typhoon track forecasts and is based on a deep sequence model (ConvLSTM) to generate corrected typhoon track forecast products. This has significantly improved the accuracy of TYTEC track forecasts.

2.1.6 Application of Global AI Weather Forecasting Models

In collaboration with Tsinghua University, the NMC has developed the "FengQing" global weather forecasting model, which is driven by the ECMWF IFS analysis fields to produce rolling forecasts (twice daily) for up to 10 days. Additionally, IFS analysis fields are used to drive models such as Pangu, Fengwu, Fuxi, GraphCast, and FourCastNet, and AIFS forecast data is also integrated to provide forecasters with comprehensive reference for weather circulations.

2.1.7 Typhoon Intensity Forecast Adaptation Model

To address the issue of underestimated typhoon intensity forecasts in global models, a CNN-based intensity forecast adaptation model has been developed. This model captures key information related to typhoon intensity by extracting features from the three-dimensional circulation field around the typhoon. Then regression to intensity is performed by dimension reduction, which adjusts the global AI model's typhoon intensity forecasts to an unbiased state and reduces the mean absolute error by 40%.

In the future, the NMC will further promote the deep integration of AI technology in typhoon monitoring, forecasting, and services, providing innovative technical support for the precise monitoring, accurate forecasting, and detailed services of typhoons worldwide. Additionally, efforts will be made to advance the development of "end-to-end" large-scale meteorological models and construct a data assimilation system based on these models to achieve satellite data assimilation to ultimately achieve a closed loop from observation to forecasting in AI global forecast models.

2.2 Hong Kong, China

HKO has been conducting real-time operational trials of various AI-based models, including Aurora, FengQing, FengWu, FuXi, GraphCast and Pangu-Weather, using ECMWF IFS operational analysis as the initial condition. Forecasts from various AI-based models are made available for the public on HKO's Earth Weather webpage (https://maps.weather.gov.hk/wxviewer/index.html?lang=en). Internal verification shows that, for many applications such as synoptic scale evolution, TC genesis and track, the AI-based models are competitive with, or even better than, conventional NWP models. HKO had also been developing and testing a fine-tuned version of Pangu based on its open-source version that could offer higher-resolution capabilities.

Starting from 2025, HKO is making operational use of real-time forecast outputs from several AI models, in addition to traditional NWP models, for the formulation of TC track forecasting using a multi-model ensemble approach. In addition to deterministic forecasts, an ensemble version of FuXi (FuXi ENS) has been trial in real time to provide probabilistic guidance on the TC track.

Leveraging lower computing costs in running AI-based models/ensembles, further technique development on applications of AI-based models/ensembles (and post-processing methods) are underway in HKO.

A major shortcoming of current AI-based models in TC forecasts is the underestimation of TC intensity. To address this, HKO adopted a hybrid technique that harnesses the advantages of both AI-based global model and high-resolution physics-based regional model. This approach uses skillful synoptic-scale forecasts from the global AI models to feed a high-resolution regional NWP model that can offer finer-scale details and better capture the intensity and wind structure of TCs. Case studies have already demonstrated the rapid intensification of Super Typhoon Yagi (2411) was well captured by the AI-assisted high-resolution regional model, with only minor short-term bias. HKO is now also referencing real-time forecasts from the high-resolution, AI-assisted regional models as part of its TC forecasting guidance.

Reference:

Lai, S. K., Y. He, P. W. Chan, B. W. Kerns, S. S. Chen, and H. Su. 2025. Towards Skillful Tropical Cyclone Forecasting by AI-Model-Driven High-Resolution Regional Coupled Model. *Meteorological Applications* 32, no. 5: e70109. https://doi.org/10.1002/met.70109.

2.3 Japan

This report presents four investigations conducted primarily by the Meteorological Research Institute of the Japan Meteorological Agency (JMA).

1. Tropical cyclone track and intensity predictions in the western North Pacific basin using Pangu-Weather and JMA initial conditions
2. Consensus forecasts of tropical cyclone track using NWP and AI models
3. High-resolution AI model for specified regions using ECMWF Anemoi framework
4. Ensemble forecasts of tropical cyclones using AI models

2.3.1 Tropical cyclone track and intensity predictions in the western North Pacific basin using Pangu-Weather and JMA initial conditions

We conducted a statistical analysis of the track and intensity forecast accuracy and characteristics of Pangu-Weather by running it with initial conditions from the Japan Meteorological Agency's (JMA) Global Spectral Model (GSM), targeting all typhoons that occurred in the western North Pacific basin between 2021 and 2023. While there have been studies on weather and tropical cyclone forecasting using AI-based weather models, few have focused specifically on typhoons in the western North Pacific and statistically evaluated the performance across many cases in comparison with numerical weather prediction (NWP) models. Our study aimed to fill this gap. One of the key findings was that the high track forecast accuracy reported in earlier studies could also be confirmed when using JMA's initial conditions. We were able to objectively examine the types of cases in which the AI model reduced forecast errors, and found that the westward (slow) bias often seen near Japan was mitigated. Another significant result was that the model maintained its forecast accuracy even for typhoons with unusual tracks. On the other hand, we also identified some limitations of the AI model. For example, it did not resolve so-called forecast bust cases—where many NWP models show large track errors—and it tended to predict weaker intensities compared to NWP models and best track data. This study has been published in the Journal of the Meteorological Society of Japan (JMSJ).

Yamagauchi, M., Y. Ikuta, K. Ito, and M. Satoh, 2025: Tropical Cyclone Track and Intensity Predictions in the Western North Pacific Basin Using Pangu-Weather and JMA Initial Conditions, Journal of the Meteorological Society of Japan. 103, 357-370.

2.3.2 Consensus forecasts of tropical cyclone track using NWP and AI models

Yamaguchi et al. (2025, JMSJ) compared typhoon track forecasts produced by the AI model Pangu-Weather with those from the JMA/GSM. This study demonstrated the potential of AI-based forecasts for tropical cyclone (TC) track prediction. In operational settings, however, consensus forecasts—typically generated by averaging outputs from multiple global models such as JMA, ECMWF, NCEP, and UKMO—are used at JMA. To evaluate the practical utility of AI models, it is therefore more relevant to compare them against these consensus forecasts rather than individual numerical weather prediction (NWP) models like the JMA/GSM. In this study, we assessed various consensus forecasts combining outputs from the four NWP models and three AI models: AIFS, GraphCast, and Pangu-Weather. The initial conditions for the AI models were derived from those of the JMA/GSM. Verification was performed for all tropical cyclones that occurred in the western North Pacific between 2021 and 2024, using forecasts initialized at 00 and 12 UTC. The results showed that combining the AI weather models with NWP models produced more accurate forecasts than consensus forecasts based solely on the AI models.

2.3.3 High-resolution AI model for specified regions using ECMWF Anemoi framework

Current AI models exhibit certain limitations:

While they predict tropical cyclone (TC) tracks well, they tend to underestimate TC intensity.

They capture the spatial distribution of precipitation relatively well but tend to underestimate its magnitude.

These shortcomings may be attributed to insufficient model resolution. To address this, we investigated whether enhancing spatial resolution in targeted regions could improve prediction accuracy. Using Anemoi, an AI-based modeling framework developed by ECMWF, we increased the resolution around Japan. The enhanced model demonstrated the capability to reproduce both extremely heavy precipitation events and very strong TC intensities.

Ikuta et al. (2025): Utilization of Machine Learning–Based Weather Prediction Models, Meteorological Society of Japan 2025 Spring Meeting.

2.3.4 Ensemble forecasts of tropical cyclones using AI models

JMA operates a global ensemble prediction system (GEPS) consisting of 51 members with a horizontal resolution of approximately 27 km. The output from this system is used, for example, to determine the size of the probability circles for typhoon track predictions. While there have been many studies evaluating the accuracy of deterministic tropical cyclone track forecasts produced by AI models, the potential of using such models within an ensemble prediction framework remains less explored. So, we used the initial conditions from GEPS as input for AI models, and assessed the characteristics of the resulting ensemble, such as the ensemble mean and spread. All 2023 TCs over the western North Pacific were included in the evaluation, with GraphCast and Pangu-Weather used as the AI models. In both GraphCast and GEPS, the ensemble mean tends to show larger track forecast errors compared to the control run. However, since the sample size may be limited, further verification will be conducted with more cases. In the track forecast experiment for Typhoon SOALA (T2309), the AI models were found to exhibit an ensemble spread similar to that of the GEPS.

Yamashita et al. (2025): Ensemble Forecasting Using AI Models Initialized with JMA Global Ensemble Prediction System, Meteorological Society of Japan 2025 Spring Meeting.

2.4 Macao, China

In Macao, China, the Meteorological and Geophysical Bureau (DSMG) is currently utilizing the ECMWF-AIFS numerical model to generate basic products for operational reference in typhoon forecasting. Additionally, the Pangu model, using ECMWF analysis fields as the initial field is under testing. The produced forecast results are temporarily for experimental purposes. Details of the model applications are provided in the accompanying table.

Model	Initial Time	Area	Spatial Resolution.	Time Resolution
Pangu	00Z, 06Z,12Z, 18Z	Global	0.25	0-360h (6h)

There are plans to adopt a broader range of AI forecast model outputs to enhance AI-driven forecasting products. Through continuous verification and analysis, these data will be progressively integrated into operational workflows to improve the accuracy and reliability of typhoon analysis and forecasting in Macao, China.

On the other hand, DSMG is also responsible for the forecasting of storm surges caused by tropical cyclones. In addition to the continued use of traditional storm surge models to forecast water level increases along the coast of the Pearl River Estuary, machine learning methods are also being used to establish a local storm surge forecasting model. This model makes predictions by establishing a relationship between local wind speed, wind direction, gusts, and air pressure with local water levels. The results have shown a certain degree of effectiveness, and DSMG is continuously verifying and optimizing this project.

2.5 Malaysia

Operational forecasters from MET Malaysia have received the AI-based TC track forecasts from the data sharing platform developed by this ETC. Visual verification of near-equatorial TCs showed promising potential in terms of track accuracy. MET Malaysia aims to perform more in-depth verification on near-equatorial TCs in addition to unnamed, weak storms that could nonetheless impact near-equatorial countries. Additionally, MET Malaysia plans to run the FourCastNet AI-model as part of its development plan towards using AI for weather forecasting.

2.6 Philippines

In the Philippines, the Department of Science and Technology - Philippine Atmospheric, Geophysical, and Astronomical Services Administration (DOST-PAGASA) has partially integrated the near-realtime TC track and intensity forecasts from the AOP's data sharing platform as reference for operations. PISTON, also called as PAGASA Integrated System for Typhoon OperatioN, serves as an integrated portal where TC track and intensity forecasts from all the models that DOST-PAGASA are receiving and it is the where operational TC forecasters are viewing, analyzing, and formulating official TC forecasts. Aside from PISTON, the data from the

AOP's data exchange platform is also integrated on an internal visualization platform for experimental systems.

In terms of running the open global AI models, DOST-PAGASA is now running GFS-driven and ECMWF-IFS-driven instances of Huawei's PanguWeather and Google DeepMind's Graphcast. Integration with a TC tracker and running more AI models is in the works. The GFS-driven and ECMWF-IFS-driven PanguWeather is initialized 6 times a day and covers a forecast time of up to 360 hrs and is encoded as a grib2 file while the GraphCast covers a forecast time of up to 144 hours. DOST-PAGASA will work on integrating that information into the data exchange server next year.

Aside from global models, DOST-PAGASA is currently testing a localized AI weather model from Atmo, Inc. and comparing it with existing NWP and AI models. Early results show that the early model versions perform slightly worse than ECMWF-HRES in terms of TC track and very poorly on TC intensity. Within the same project, in collaboration with DOST-ASTI (Advanced Science & Technology Institute), a localized AI weather model is planned to be done using ECMWF Anemoi and PAGASA-WRF data.

2.7 Republic of Korea

The Korea Meteorological Administration (KMA) has operationalized an AI-based global weather forecasting system to provide forecasts with new insights and to evaluate the performance and applicability of AI-based weather forecast models since February 2024. The AI-based global forecasting system consists of nine model configurations, combining three AI-based weather forecasting models (FourCastNet2 from NVIDIA, Pangu-Weather from Huawei, and GraphCast from Google DeepMind) with three initial conditions which are produced by traditional numerical weather prediction models, such as KIM (Korea Integrated Model), UM (Unified Model) and ECMWF IFS (Integrated Forecast System). Each configuration shares common elements, except AI-model and Initial condition used: 6-hourly forecasts, 0.25° spatial resolution, and 13 vertical pressure levels, and each provides forecasts extending up to +288 hours (Table 2.1.). Forecasts are produced twice daily using the 00 and 12 UTC initial conditions. Each model outputs surface and upper-air fields, stored in NetCDF format, with over 20 meteorological variables including u, v, T, q, hgt, psl, t2m, u10m, and v10m. These outputs are used as supplementary guidance for synoptic-scale forecasting, including precipitation distribution, temperature trends, and pressure field evolution. The AI models offer an alternative perspective to complement traditional NWP guidance, enhancing forecasters' situational awareness and supporting more informed operational decisions.

Table 2.1. Summary of AI weather forecasting models in KMA.

	Pangu-Weather	GraphCast	FourCastNet2
Developer	Huawei	Google DeepMind	NVIDIA
Initial Conditions Used	KIM/UM/IFS	KIM/UM/IFS	KIM/UM/IFS
Forecast Range	Up to +288 h	Up to +288 h	Up to +288 h
Time Step	6-hours intervals	6-hours intervals	6-hours intervals
Horizontal Resolution	25 km	25 km	25 km
Vertical Resolution	13 layers	13 layers	13 layers
Run Schedule	Every 2 hours (00, 12 UTC)	Every 2 hours (00, 12 UTC)	Every 2 hours (00, 12 UTC)
Training Data	ERA5	Pre-training: ERA5(1979-2017) Fine-tuning: ECMWF HRES-fc0(2016-2021)	ERA5

For typhoon-specific forecasting, KMA applies the GFDL Vortex Tracker to the outputs from all nine AI prediction configurations (three models × three initial conditions). This allows real-time extraction of typhoon center location, intensity (central pressure and maximum wind speed), and wind radii throughout the forecast range. The outputs are incorporated into KMA's Typhoon Operation System, where they are visualized alongside traditional NWP forecasts. This setup enables forecasters to directly compare multiple guidance sources and refine official forecasts in real time.

KMA is advancing its next-generation AI-based forecasting capabilities through the integration of new AI-models and development of in-house models. Newly developed AI-based weather forecasting model, such as AIFS from ECMWF and GenCast from Google DeepMine, is currently being incorporated into operations. In addition, KMA has launched projects to develop in-house AI-based weather forecasting models optimized for East Asia. A foundation Model for Weather and Climate, such as Aurora from Microsoft, is planned to be developed by 2028 to support multi-scale spatiotemporal forecasting. Along with the global weather forecasting model, KMA is also focusing on improving AI-based nowcasting model and applying xAI (explainable AI) methods on it.

3. Typhoon track and intensity predictions by AI models for 2024 and 2025 typhoons

3.1 Data and methodology

During the 2024 Western North Pacific typhoon season, a total of 19 AI-based tropical cyclone (TC) forecasts were submitted by CMA, HKO, JMA, KMA, and SMS (Shanghai Meteorological Service, whose data is provided via STI (Shanghai Typhoon Institute)) by contributing Members of the Typhoon Committee ET-AITC. The contributing models varied in architecture and operational settings (Table 3.1.1). CMA and SMS, both representing China, submitted forecasts using the FGQG_IFS, FENW_IFS, and FUXI_IFS models. HKO contributed four AI-WFMs (FENW, FUXI, GRAP, and PANG) all initialized with IFS. JMA contributed models AIFS, GRAP, and PANG with GSM initialization. KMA submitted nine combinations of PANG, GRAP, and FOUR models using three ICs: IFS, KIM, and UM. Most models generated forecasts up to 120 hours from 0000 and 1200 UTC cycles. However, HKO issued forecasts four times daily (0000, 0600, 1200, and 1800 UTC), while JMA provided the shortest forecast range and KMA extended forecasts up to 288 hours. Most models were trained using ERA5 reanalysis. For typhoon tracking, KMA and HKO applied the GFDL Vortex Tracker algorithm to extract center position, intensity, and wind radii from each AI forecast field. CMA used the minimum of sea-level pressure to determine the typhoon center. SMS used the minimum of sea-level pressure, vorticity, and the vertical structure of the atmosphere to determine the typhoon center. JMA also used the minimum sea-level pressure for TC tracking (Yamaguchi et al., 2025).

Table 3.1.1. Summary of AI-based tropical cyclone (TC) forecast models evaluated in this study.

Member	Agency	# of model	AI Model name	Initial Data	Abbreviation	Real-time Running	Forecast length (h)	Times of daily run (UTC)	Tracking Algorithm
China	CMA	1	Fengqing	IFS	CMA_FENQ_IFS	Y	240	00, 12	Min. SLP
	SMS	2	Fengwu	IFS	SMS_FENW_IFS	Y	168	00, 12	Min. SLP, VOR, Vertical Structure of the Atmosphere
			Fuxi		SMS_FUXI_IFS				
Hong Kong, China	HKO	4	Fengwu	IFS	HKO_FENW_IFS	Y	240	00, 06, 12, 18	GFDL Vortex Tracker
			Fuxi		HKO_FUXI_IFS				
			GraphCast		HKO_GRAP_IFS				
			Pangu-Weather		HKO_PANG_IFS				
Japan	JMA	3	Pangu-Weather	GSM	JMA_PANG_GSM	N	132	00, 12	Min. SLP
			AIFS		JMA_AIFS_GSM				
			GraphCast		JMA_GRAP_GSM				
Republic of Korea	KMA	9	Pangu-Weather	IFS	KMA_PANG_IFS	Y	288	00, 12	GFDL Vortex Tracker
				KIM	KMA_PANG_KIM				
				UM	KMA_PANG_UMM				
			GraphCast	IFS	KMA_GRAP_IFS				
				KIM	KMA_GRAP_KIM				
				UM	KMA_GRAP_UMM				
			FourCastNet	IFS	KMA_FOUR_IFS				
				KIM	KMA_FOUR_KIM				
				UM	KMA_FOUR_UMM				

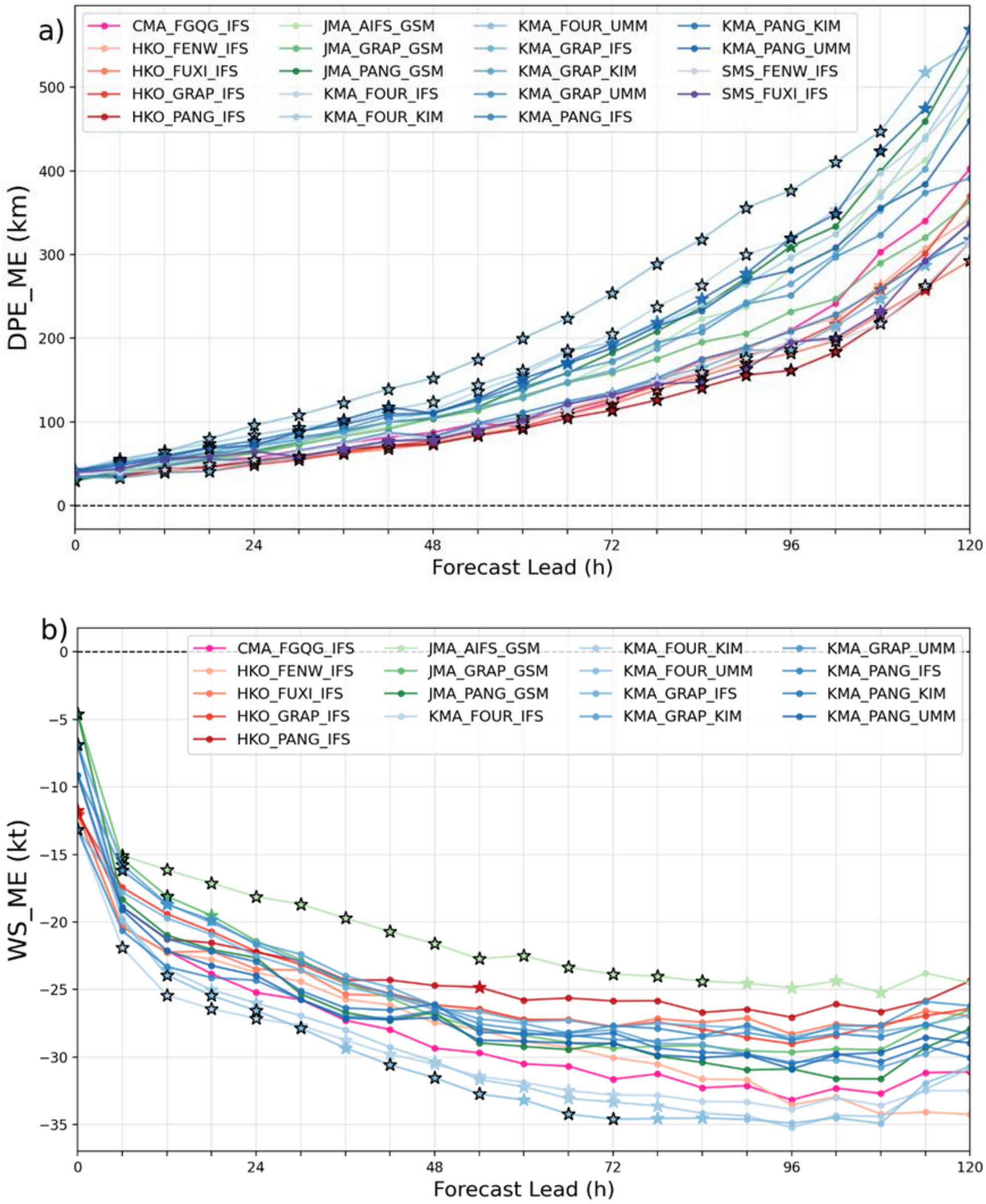
3.2 Verification results

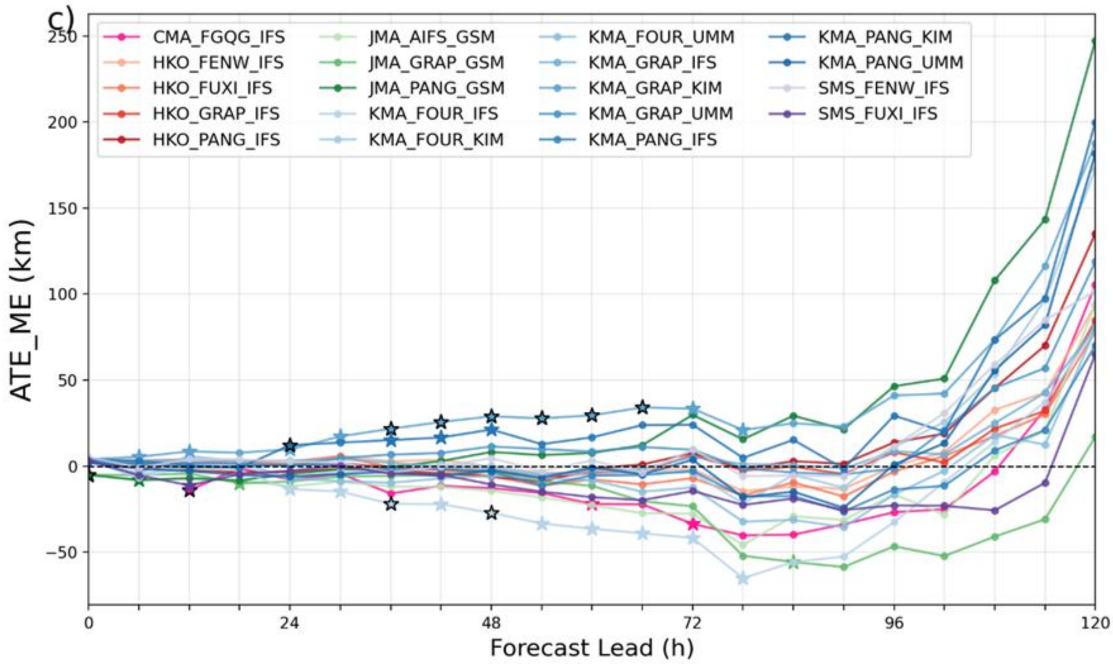
A verification of 19 AI-based TC forecasts was conducted for the 2024 Western North Pacific typhoon season using best-track data provided by RSMC Tokyo. While track forecasts from all 19 models were evaluated, only 17 models were included in the intensity verification because two

SMS submissions lacked sufficient wind information. Specifically, SMS_FENW_IFS contained too few valid maximum wind speed (MWS) samples, and SMS_FUXI_IFS did not provide MWS output. In addition, CMA forecasts were excluded from the 0- and 6-hour samples for both track and intensity, due to missing data at those initial lead times. Forecast performance was assessed up to 120 hours using standard metrics: direct position error (DPE) for track accuracy, wind speed mean error (WS_ME) for intensity, and positional vector decomposition (ATE, CTE).

Most models exhibited increasing DPE with lead time, but several showed comparatively high skill throughout the forecast range. In particular, HKO_FENW_IFS, HKO_FUXI_IFS, KMA_GRAP_IFS, SMS_FENW_IFS, and SMS_FUXI_IFS consistently reported smaller track errors than other models (Figure 3.2.1a). For intensity forecasts, all models tended to underestimate the maximum sustained wind speed. JMA_AIFS_GSM showed the least negative bias, and this difference was statistically significant compared with most other models, particularly within 84 hours (Figure 3.2.1b). For ATE, most models showed near-zero values up to 24 hours, indicating accurate translation speed prediction in the early forecast range. However, by 120 hours all models produced positive ATEs, reflecting a general tendency to overestimate typhoon forward speed. Notably, KMA_GRAP_KIM showed a statistically significant faster motion than observed during the 24–66 h range (Figure 3.2.1c). For CTE, most models initially exhibited a slight left-of-track bias, which gradually shifted to a right-of-track displacement with longer lead times. This indicates a systematic right bias at extended ranges. In particular, HKO_GRAP_IFS, KMA_FOUR_IFS, and KMA_GRAP_IFS displayed significantly positive mid-range CTEs, confirming a persistent rightward bias. By contrast, JMA_PANG_GSM and KMA_PANG_UMM maintained statistically significant leftward biases during the 24–54 h period (Figure 3.2.1d).

To evaluate the practical utility of AI-based typhoon forecasts, their performance was compared against official operational forecasts and traditional NWP models. Figure 3.2.2 presents a homogeneous comparison of forecast performance in terms of mean direct position error (DPE_ME) for track prediction and mean bias error of maximum wind speed (WS_ME) for intensity. For this intercomparison, representative AI-based models were selected separately for track and intensity, based on the best-performing model from each contributing Member. Specifically, HKO_PANG_IFS, JMA_GRAP_GSM, KMA_GRAP_IFS, and SMS_FENW_IFS were chosen for track (DPE_ME), while HKO_PANG_IFS, JMA_AIFS_GSM, KMA_GRAP_KIM, and CMA_FQGQ_IFS were selected for intensity (WS_ME). The official forecast group comprised guidance issued by CMA, HKO, JMA (RSMC Tokyo), JTWC, and KMA, while the traditional NWP group included the ECMWF IFS and JMA GSM. Results indicate that for track forecasts (Fig. 3.2.2a), AI-based models consistently produced the smallest mean position errors, with the advantage becoming increasingly pronounced at longer lead times. In contrast, official forecasts and NWP models exhibited nearly identical error growth, reflecting their operational interdependence. For intensity forecasts (Fig. 3.2.2b), AI models systematically underestimated TC intensity, showing larger negative wind speed biases at all lead times. Official forecasts and NWP models, by comparison, maintained smaller-magnitude biases, often more than 50% lower than those of AI systems. These findings highlight a marked contrast: AI-based models already provide superior long-range skill in TC track prediction but remain limited in accurately capturing TC intensity.





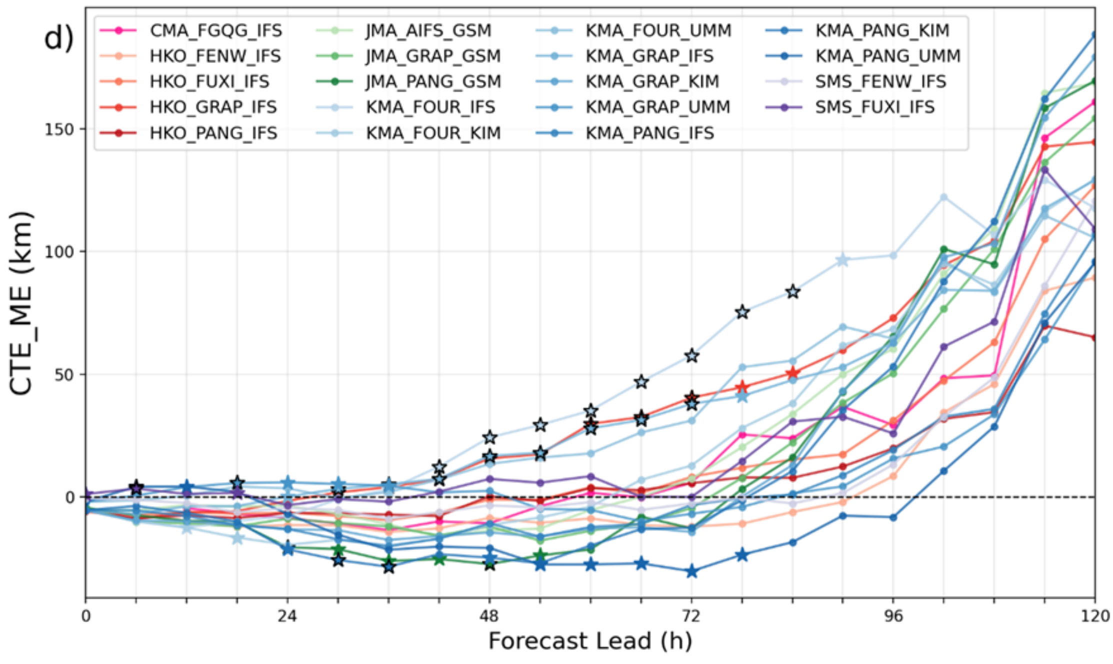


Figure 3.2.1. Mean forecast errors for the 19 AI-TC forecast models over the 2024 WNP season as a function of forecast lead time: (a) direct position error (DPE_ME, km), (b) maximum wind speed bias (WS_ME, kt), (c) along-track error (ATE_ME, km), and (d) cross-track error (CTE_ME, km). Colored stars denote differences between an individual model and the multi-model mean that are statistically significant at the 95% confidence level; black-edged bold stars indicate significance at the 99% level (two-sided Welch's *t*-test).

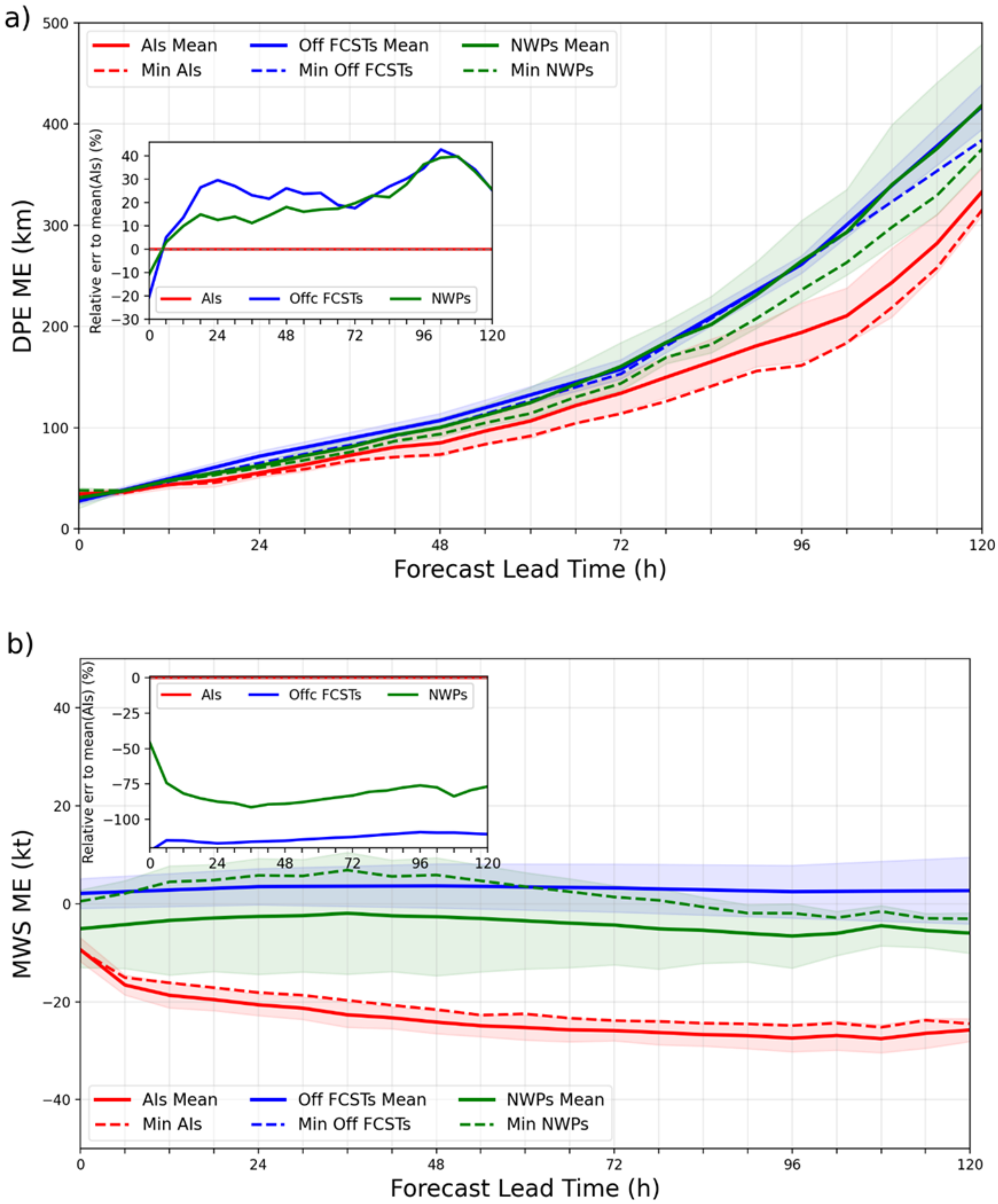


Figure 3.2.2. Homogeneous comparison of forecast performance in terms of (a) mean direct position error (DPE_ME, km) and (b) mean bias error of maximum wind speed (WS_ME, kt) across different forecast lead times. Three forecast system groups are evaluated: representative AI-based models (red), official forecasts issued by national meteorological agencies (blue), and traditional numerical weather prediction (NWP) models (green). The AI group comprises the best-performing model for each contributing Members and verification variable (DPE_ME and WS_ME): HKO_PANG_IFS, JMA_GRAP_GSM, KMA_GRAP_IFS, and SMS_FENW_IFS for

DPE_ME, and HKO_PANG_IFS, JMA_AIFS_GSM, KMA_GRAP_KIM, and CMA_FQGQ_IFS for WS_ME. Official forecasts include CMA, HKO, JMA, JTWC, and KMA, while NWP systems include IFS and GSM. Solid lines represent the group means, shaded areas indicate one standard deviation, and dashed lines denote the best-performing model within each group. The inset panels in both (a) and (b) present relative errors (%) concerning the AI ensemble mean, computed as (GroupMean−AIMean)/AIMean×100.

3.3 Discussion

The intercomparison exercise demonstrates that AI-based tropical cyclone (TC) forecasting has reached a stage where it can provide meaningful guidance for operational use, particularly in track prediction. At the same time, the results highlight clear limitations in intensity forecasting and reveal differences in performance that stem not only from the AI model architecture itself but also from the quality of the initial conditions used to drive the forecasts (Table S1-3). The evaluation further revealed that, even when models employed the same AI framework and initial conditions, their TC forecast performance could differ substantially (Table S4, Figure S2). This suggests that factors such as implementation details, input data preprocessing, and storm tracking algorithms can critically shape model performance.

Despite these strengths, several limitations remain. Most AI systems systematically underestimate typhoon intensity, particularly failing to capture episodes of rapid intensification, while showing some ability to reproduce weakening phases (Figure S1). In addition, since this assessment covered only a single season, a broader evaluation using multiple years and larger storm samples is needed to ensure robustness under diverse conditions. The analysis also indicated that along-track (ATE) and cross-track (CTE) errors grew substantially at longer lead times, a result strongly influenced by large errors during Typhoon Shanshan in 2024. Such case sensitivity emphasizes the importance of multi-season verification to prevent results from being dominated by individual storms.

Further efforts may usefully explore the development of high-resolution AI forecasting systems, AI-based Earth system prediction platforms, hybrid methods combining AI and high-resolution NWP, and ensemble-based approaches. Equally important will be the systematic sharing of datasets, verification protocols, and operational experiences across agencies. Continued international collaboration under the Typhoon Committee ET-AITC framework will therefore be critical not only for advancing technical development but also for accelerating the integration of AI guidance into operational TC forecasting practices in the region.

3.4 User experience and good practices

A case study on Severe Typhoon Danas (2025)

Severe Typhoon Danas over the northern part of the South China Sea and seas near Taiwan in early July 2025 had an erratic path. It eventually turned at right angle or acute angle twice over the northeastern part of the South China Sea and the East China Sea. Danas intensified into a tropical storm and was named in the small hours of 5 July. The sequence of pre-genesis and early-stage TC forecast tracks by NWP models and AI models based on 12 UTC forecasts of 30 June up to 4 July 2025 is shown in Figure 3.4.1. As early as 30 June 2025 (Figure 3.4.1a), majority of the models had indicated TC genesis over the region. At that time, NWP models suggest a TC may form and move across the western north Pacific; while AI models suggest that the TC, once formed, is expected to enter the South China Sea, which was much closer to the later evolution. However, AI models by then gave forecast tracks generally towards the western coast of Guangdong, with considerable error. The following day (Figure 3.4.1b), majority of the AI models started to get it right for the 90-degree track change over the northeastern part of the South China Sea for this TC towards Taiwan, whereas NWP models were still forecasting the scenario of moving across the western North Pacific. The NWP models began to become more consistent with the actual track on 2 and 3 July 2025 (Figure 3.4.1c,d). Eventually, all models aligned in their forecasts of 4 July (Figure 3.4.1e).

In this case, the genesis and movement of Danas were not captured well by the conventional physics-based global NWP models in their medium-range forecasts, while AI models were able to suggest more accurately the genesis position and its earlier movement, highlighting AI models' value in particular for forecasts of longer lead times.

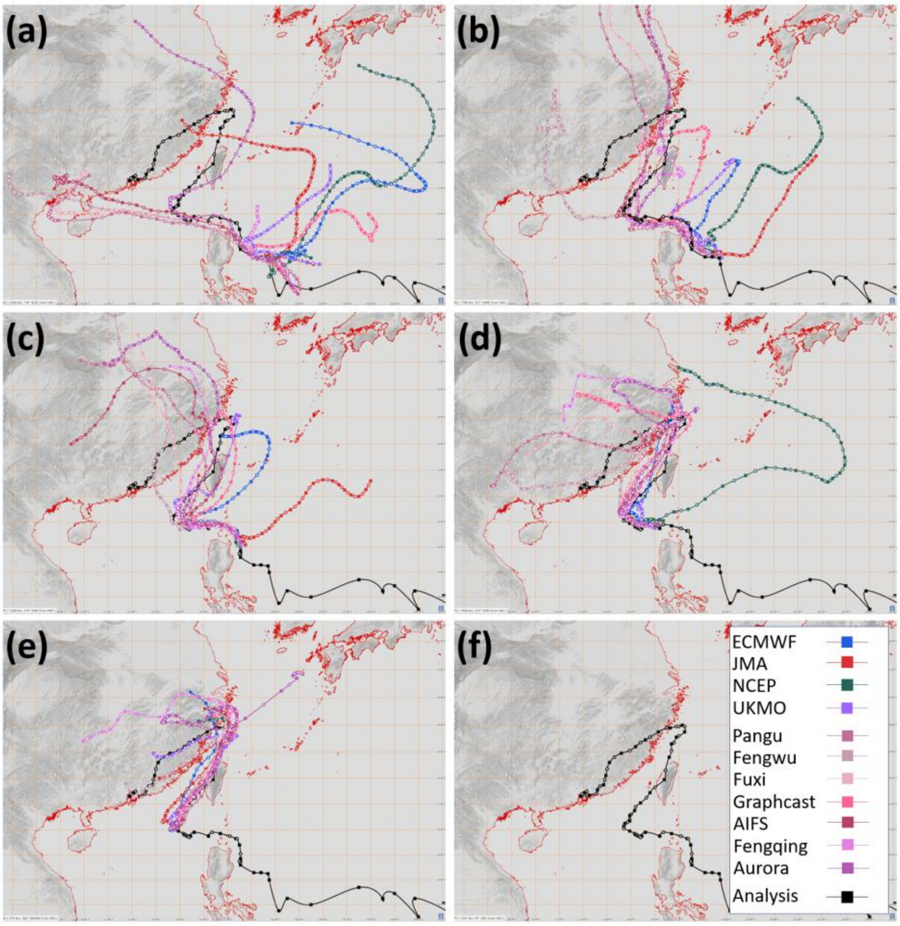


Figure 3.4.1. Pre-genesis and early-stage TC track forecast by NWP models (ECMWF, JMA, NCEP and UKMO) and AI models (Pangu, Fengwu, Fuxi, Graphcast, AIFS, Fengqing and Aurora) for TC Danas. Model runs are 12 UTC forecast on (a) 30 June 2025; (b) 1 July 2025; (c) 2 July 2025; (d) 3 July 2025, and (e) 4 July 2025. (f) shows the operational analysis positions, with both pre-genesis stage of the TC and its remnant after landfall. Figure provided by HKO.

4. Data exchange and policy

4.1 Establishment of data exchange platform

Hong Kong Observatory of Hong Kong, China contributed to this AOP by developing and operating a Data Exchange Portal on behalf of Typhoon Committee. The Portal is a closed web platform for data exchange among TC Members, accessible from https://aiproject.typhooncommittee.org/. It provides the necessary infrastructure for real-time, near-real-time and none-real-time data exchange (as various contributing Members would choose), as well as archival of TC forecast tracks from a number of AI models.

4.2 Data shared

As of 2025, tropical cyclone forecast tracks as extracted from the output of various AI models for named tropical cyclones (as identified by RSMC Tokyo). The following fields are included:

- Forecast lead time (hour)
- Latitude (degree N)
- Longitude (degree E)
- MSLP (hPa)
- Maximum sustained wind speed (knot)

The AI models contributed by members are listed in Section 3.

4.3 Data exchange policy

Noting that the primary objective of this AOP is to share findings, including verification results and good practices, and to explore how AI-based tropical cyclone analysis and forecasting can be effectively utilized in the future, the Expert Team considered that it is important to be mindful on the use of data being exchanged. The Expert Team also noted that licensing requirements of many AI models adopt public copyright licenses (such as Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)) that enable the free distribution of the model output. The following data exchange policy is agreed with the Expert Team.

- Data will be exchanged among Typhoon Committee Members with representatives in the Expert Team.
- Data exchanged for this AOP will not be provided to third parties.
- The data will only be used for the purposes of this AOP and will not be utilized for any purposes beyond the scope of this AOP. For example, data exchange arrangement is intended to support the research activities under this AOP and not for operational tropical cyclone forecasting.
- Acknowledgement will be given on using the data from this AOP.

The Expert Team will review the data exchange policy in 2026.

5. Meetings

5.1 Expert team in-person meeting in June 2025 (Tokyo, Japan)

From June 24 (Tuesday) to June 26 (Thursday), 2025, a meeting of the Expert Team on AI Applications in Tropical Cyclone Analysis and Forecasting (ET-AITC), under the Working Group on Meteorology of the Typhoon Committee of the United Nations Economic and Social Commission for Asia and the Pacific (ESCAP) and the World Meteorological Organization (WMO), was held at the Japan Meteorological Agency (JMA) in a hybrid format, with both in-person and online participation. Program of the Expert Team meeting is in Appendix 2.

On the first day, members of the expert team reported on recent developments related to AI technologies for tropical cyclone analysis and forecasting, including their utilization at operation, in a format that was also made accessible online to Members of the Typhoon Committee. In addition, invited lectures were delivered by developers of AI weather models, presenting the latest advancements in cutting-edge systems.

On the morning of the second day, the Korea Meteorological Administration (KMA) presented interim results of verification work on tropical cyclone track and intensity forecasts produced by AI weather models for typhoons in 2024 (details are provided in Section 3). Despite challenges such as incomplete or inconsistent data and a very limited time of less than one month for the verification work, KMA conducted thorough analysis and presented highly detailed results. The expert team commended KMA's dedication and effort.

In the afternoon of the second day, a presentation was given on the establishment and usage of an online platform for data exchange. This was followed by discussions on data exchange policies. The team also discussed the structure and drafting plan of the project's annual activity report, which is scheduled to be completed by the IWS in November 2025.

On the third day, discussions were held on the future activity schedule, upcoming workshop plans, collaboration with the WMO Integrated Processing and Prediction System (WIPPS) and the WMO Advisory Group on Tropical Cyclones (AG-TC), and further user requirements. It was confirmed that regular online meetings would continue going forward. Participants also proposed the possibility of holding in-person meetings and workshops in Shanghai, China in spring 2026 and in the Philippines in 2027. Coordination efforts will begin in preparation for the potential meeting in spring 2026.

For 2026, it was agreed that intercomparison of AI-based tropical cyclone forecasts would continue as a core activity. In addition, new proposals were made for intercomparison of AI applications in tropical cyclone analysis and for verification focusing on tropical cyclone genesis forecasts.

Group photo of the participants of the expert team in-person meeting in June 2025 (Tokyo, Japan)

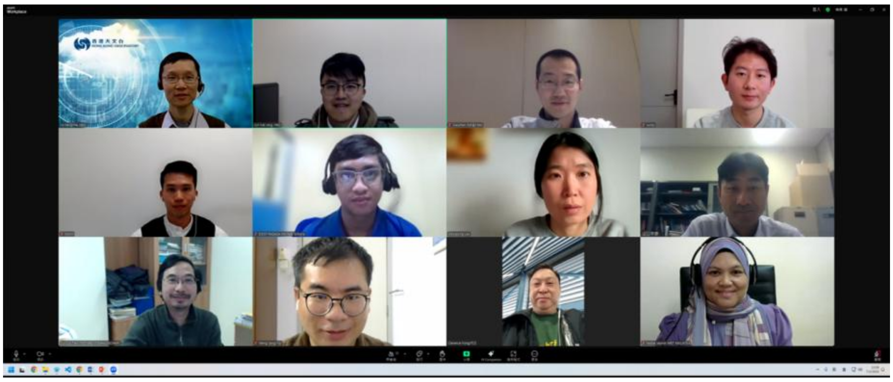
5.2 Online meetings

A total of four online meetings have been held to date. Anticipating the formal approval of the project at the annual session in March 2025, the first online meeting was organized in February 2025 as part of the preparation process.

Prior to the first online meeting, a survey was conducted among prospective members of the expert team to understand how they could contribute to the project. Based on the survey results, the first online meeting focused on discussions regarding the planned activities for 2025. It was agreed that the primary activity would be an intercomparison of typhoon forecasts using AI-based weather models. In addition, discussions on holding an in-person meeting of the expert team were initiated.

In the subsequent second online meeting, it was confirmed that CMA/NMC, CMA/STI, HKO, and JMA would provide typhoon forecast data generated by AI weather models. Later, KMA also joined as a data provider. It was also agreed that KMA would take the lead in verification tasks. Furthermore, HKO and TSC would jointly lead the establishment of a data exchange platform, and the in-person expert team meeting would be hosted by JMA in June 2025.

During the third online meeting, detailed discussions were held on file naming conventions and data formats for typhoon forecast data produced by AI weather models. A common specification was agreed upon, and it was decided that the forecast data for all typhoons occurring in 2024 would be submitted by May 24, 2025. HKO reported that preparations for the data exchange platform were progressing as scheduled, and JMA shared logistical information regarding the in-person meeting.



Group photo of the first online meeting of ET-AITC

Online Meeting	Agenda
2025.02.07	1. Round-table introduction 2. Background of the project and work plan for 2025 3. Potential contribution from each Member 4. Possible in-person meeting of the Expert Team and/or a workshop 5. AOB
2025.03.21	1. Contribution from each Member 2. Discussion on TC track data exchange 3. Discussion on data policy 4. Requirements on data exchange platform 5. In-person meeting and travel support 6. AOB
2025.04.28	1. TC track data exchange 2. Technical details on verification 3. Progress on data exchange platform 4. ET on-site meeting in June in Tokyo 5. AOB

2025.08.28	1. Follow-up on action items from the Tokyo meeting in June 2. Workshop in Shanghai scheduled for April 2026 3. Proposed budget for the 2026 activities 4. Discussion on activities for 2026 5. AOB

6. Future plan

6.1 Further user requirement

Artificial intelligence-based tools such as AI weather models are deemed to be helpful as an additional tool to the operational forecasters as demonstrated on the intercomparison. However, there is still some resistance to adopting its use in operational settings. Explainability and performance are some of the main reasons for the hesitance of operational forecasters. To address that, it is recommended to examine individual cases in which AI-based models had good and bad performance on its TC forecast.

Near the equator, Members like Malaysia and the Philippines suffer from effects of low-pressure systems and TC-like vortices prior to it becoming a TC. These cases become prevalent on the Philippine area during the early 2025 typhoon season. As a test of the operational use of AI models, operational forecasters from PAGASA checked the near real time AI based TC track forecasts from HKO on the data exchange server and found out that the TC Forecasts on the data exchange platform appears 6-18 hours after the initial official forecast have been issued. The TC forecast data from the exchange platform starts when the TC has been named by the Tokyo-RSMC (e.g. it reached maximum sustained winds of 34 kts and above - Tropical Storm category). Meanwhile, PAGASA starts issuing official forecasts on the LPA stage as long as it is within the PAGASA Tropical Cyclone Information Domain (TCID). The differences in the stage of TC development when the system was started to track strengthens the need to also track weak systems.

Another aspect of this AOP that have not been tackled as a whole in this year is the tropical cyclone analysis. Several agencies such as KMA, HKO, and CMA have developed their own AI based TC monitoring systems using satellite data. An intercomparison study is recommended to be done.

In terms of developing the capacity on AI model development and application, it is good to look at how some Members in the region, such as HKO, KMA, and CMA, have successfully created, used, and integrated AI tools in their operational TC forecasting and analysis.

6.2 Potential collaboration with other WMO programs and initiatives

The ninth meeting of the Advisory Group on Tropical Cyclones (AG-TC) was held at the Japan Meteorological Agency (JMA) from May 20 to 22, 2025, where the activities of this AOP were presented. AG-TC is composed of designated representative of TC RSMCs and of TCWCs established by a WMO constituent body and sits under the Standing Committee on Disaster Risk Reduction and Public Services (SC-DRR) of the Technical Commission for Services (SERCOM), WMO.

At the 10th International Workshop on Tropical Cyclones held in Bali, Indonesia, in 2022, one of the key recommendations was as follows, highlighted the importance of integrated research and operations: "*Recommend further research into explainable and validated AI/ML techniques, in cooperation with the operational community, to address components in the tropical cyclone analysis and forecasting process.*"

Against this backdrop, the AG-TC has shown strong interest in AI-related projects such as this AOP and requested an introduction to this AOP. Mr. Yamaguchi presented the launch and objectives of the AOP. The AG-TC has also requested that progress on the AOP's activities be reported at the next AG-TC meeting, which will be held online in December 2025.

One advantage of this collaboration between the AOP and AG-TC is that it provides an opportunity to understand how AI technologies are being applied at operational centers such as RSMCs and TCWCs. It also facilitates gathering information on the needs of these centers, and by sharing such insights with ET members, it is expected to contribute to further development of the AOP.

The WMO Integrated Processing and Prediction System (WIPPS) pilot projects (https://community.wmo.int/en/wipps-pilot-project) aim to advance WIPPS by developing and adopting new technologies based on the projects' outcomes. These projects are designed through a co-production approach with WMO Members and partner organizations. The goal is to enhance the accessibility and quality of WIPPS products for all WMO Members, thereby strengthening forecasting capabilities across timescales and disciplines, including impact-based forecasting, multi-hazard early warning services, disaster risk reduction and climate change adaptation. To increase the visibility of this AOP, share success stories, and potentially engage a boarder community to advance tropical cyclone analysis and forecasting, it is proposed to join the WIPPS Pilot Project.

6.3 Next workshop and meetings

Mr Sun, representative from STI, has proposed holding a workshop and Expert Team meeting in Shanghai in April 2026. It was decided to coordinate with the co-chairs, the TCS, and relevant parties in Shanghai to plan and prepare the event. Additionally, Mr Simora, Expert from the Philippines, reported the possibility of hosting a workshop and Expert Team meeting in the Philippines in future (e.g. 2027). The Expert Team very much appreciated contributions from Members on continuation of a series of Workshops and meetings.

6.4 Work plan for 2026

Target activities in 2026 are listed below:

- Continue TC forecast verification for 2025. CMA, STI, JMA, and KMA had confirmed that TC forecasts up to a lead time of 240 hours would be provided starting in 2025, on a non-real-time basis. While HKO had already been providing forecasts up to 240 hours ahead through the data exchange platform in real time.
- Inter-comparison of TC analysis using satellite-based AI estimates for TC location and intensity. CMA, STI, HKO, and KMA confirmed that their data could be provided for the inter-comparison of typhoon analysis. STI will take the lead of verification work.
- Verification for weak TC or tropical disturbance (e.g. Tropical Depression, Tropical Storm, Tropical Low). This is to be led by Malaysia.
- To consider verification of pre-genesis tracks, pending further discussion on the tracking and identification of potential TCs/lows.
- Verification of TC genesis (e.g. how many days prior to formation have the TC been forecasted by AI models). This would be led by HKO and Vietnam.
- Individual case studies of individual TCs, to be led by Macao, China.

7. Summary and Recommendation

This AOP has been producing very fruitful results in 2025 through meetings, discussions and operation of the data exchange platform. It is recommended that this AOP will be continued in 2026 and will expand its activities as detailed in Section 6 above. Special funding will also be requested to support TC Members' participation in the Workshop in April 2026 (Section 6.3).

Appendix 1 Members of the Expert Team

Members of the Expert Team (as of September 2025)

Representative	Typhoon Committee Member
Gaozhen NIE	China (CMA)
Ziyao SUN	China (STI)
Yu-heng HE, co-chair	Hong Kong, China
Yuk-Sing LUI	Hong Kong, China
Munehiko YAMAGUCHI, co-chair	Japan
Kuok Hou HO	Macao, China
Abdul Aizat Nazmi A AZMI	Malaysia
Weng Sang YIP	Malaysia
Michael SIMORA	the Philippines
Woojeong LEE	Republic of Korea
Du Duc TIEN	Socialist Republic of Viet Nam
Eric LAU	United States of America
Clarence FONG	Typhoon Committee Secretariat

Appendix 2 First face-to-face meeting of ET-AITC

24 -26 June 2025, Tokyo, Japan

Tentative Work Plan

Time	24 June (Tue)	25 June (Wed)	26 June (Thu)
09:00 – 12:30	• Opening (9:00-9:30) ○ Welcome [DG of Atmosphere and Ocean Department, JMA] ○ Opening remark by TCS [Mr Fong] ○ Self-introduction by ET Members ○ Group photo • Presentations on latest technological advancements by ET Members (9:40-12:30) ○ HKO [Mr He] ○ CMA/NMC [Mr Nie] ○ CMA/STI [Ms Yang on behalf of Mr Sun] ○ MET Malaysia [Mr Sang] ○ KMA [Ms Lee] ○ JMA [Mr Yamaguchi]	• Adoption of agenda and review of actions (9:00-9:30) [Led by Mr Yamaguchi] • Summary of verification results in past TC cases (9:30-10:30) [Led by Ms Lee] • Discussion about verification results (10:50-11:30) • Review of data exchange and further requirements (including data policy) (11:30-12:30) [Led by Mr He]	• Discussion on future user requirements (9:00-9:50) [Led by Mr Yip, Mr Simora and Mr Du] • Discussion on work plan 2026-2027 (9:50-10:40) • Potential collaboration with other WMO programs and initiatives (11:00-11:30) [Led by Mr Yamaguchi and He] • Next workshop and meetings (11:30-12:10) • AOB (12:10-12:30)
	Lunch		

14:00 – 17:30	• Invited talks from AI development teams 14:00-14:15 (0500-0515UTC) [JMA, Mr Wakamatsu] 14:15-15:00 (0515-0600UTC) [Fengwu, Mr Han] 15:00-15:45 (0600-0645UTC) [Univ. Cambridge, Ms Allen] 15:45-16:00 (0645-0700UTC) break 16:00-16:45 (0700-0745UTC) [AIFS, Mr Maier-Gerber] (Cancelled) 16:45:17:30 (0745-0830UTC) [Pangu, Mr Bi]	• Visit to JMA operational room (14:00-14:50) • Discussion toward compilation of verification report, good practice and guidelines (15:00-17:30) ○ Structure of the report ○ Additional work that will be needed in the future to prepare the report ○ Assignment of responsibilities to each member of the team	
	Reception dinner	/	/

List of participants

Expert Team Members	
Name	**Affiliation**
Mr. Gaozhen NIE	National Meteorological Center, China Meteorological Administration
Mr. Ziyao SUN	Shanghai Typhoon Institute, China Meteorological Administration ※
Mr. Yu-Heng HE (Co-chair)	Hong Kong Observatory
Mr. Yuk-Sing LUI	Hong Kong Observatory ※
Mr. Munehiko YAMAGUCHI (Co-chair)	Japan Meteorological Agency
Ms. Woojeong LEE	Korea Meteorological Administration
Mr. Kuok Hou HO	Macao Meteorological and Geophysical Bureau ※
Mr. Abdul Aizat Nazmi BIN A AZMI	Malaysian Meteorological Department ※
Mr. Weng Sang YIP	Malaysian Meteorological Department
Mr. Michael SIMORA	Philippine Atmospheric, Geophysical and Astronomical Services Administration
Mr. DU DUC Tien	National Center For Hydro-Meteorological Forcasting, Vietnam
Mr. Clarence FONG	Typhoon Committee Secretariat
In-person Attendance by Proxy	
Ms. Mengqi YANG	Shanghai Typhoon Institute, China Meteorological Administration ※
Mr. Chi Hang PUN	Macao Meteorological and Geophysical Bureau
Invited Speakers	
Mr. Shunya WAKAMATSU	Japan Meteorological Agency
Mr. Tao Han	Shanghai Artificial Intelligence Laboratory ※
Ms. Anna Allen	University of Cambridge ※
Mr. Kaifeng Bi	Huawei ※

(※ Online participants)

Table S1. Mean track (km) and intensity forecast errors (kt) for the 19 AI models averaged for the 24-, 48-, 72-, 96- and 120 h forecasts for the 2024 in the NWP. Results are based on 19 models for DPE_ME and 17 models for MWS_ME. Mean and standard deviation for each forecast time are listed in the last row, respectively. Statistical significance is estimated based upon a two-sided Welch's *t* test. One / two asterisk indicate that the difference between an individual model and corresponding ensemble mean (with the model included) is significant at a 95% / 99% confidence level.

		DPE_ME					MWS_ME				
		Forecast lead times					Forecast lead times				
Organization	AI models	24	48	72	96	120	24	48	72	96	120
CMA	FGQG_IFS	56.6	87.3	123.9*	209.7	402.5	-25.2	-29.3	-31.6	-33.2	-31.1
HKO	FENW_IFS	52.0**	75.0**	130.3**	209.3	343.3	-23.7	-27.4	-30.0	-33.5	-34.2
	FUXI_IFS	53.8**	73.3**	121.3**	182.0**	292.8**	-23.5	-26.5	-27.7	-28.3	-26.9
	GRAP_IFS	48.8**	77.4**	127.1**	191.7**	369.6	-22.2	-26.1	-27.8	-29.0	-26.4
	PANG_IFS	53.5**	73.6**	113.9**	161.5**	315.2	-22.2	-24.7	-25.8	-27.0	-24.4
JMA	AIFS_GSM	63.3	105.1	170.7	283.1	479.1	-18.1**	-21.6**	-23.8**	-24.9*	-24.5
	GRAP_GSM	62.2	104.3	158.6	231.8	363.5	-21.4	-27.0	-29.3	-29.6	-26.5
	PANG_GSM	64.9	104.1	182.7	309.8*	553.3	-22.6	-26.7	-29.0	-30.9	-27.9
KMA	FOUR_IFS	68.2	113.8*	204.8**	317.2*	495.9	-27.1**	-30.5*	-32.8*	-33.9	-32.5
	FOUR_KIM	84.3**	124.0**	194.5*	296.6	520.3	-26.0*	-30.3*	-33.3*	-35.2	-30.8

	FOUR_UMM	96.0**	152.2**	253.7**	376.6**	551.6	-26.6**	-31.5**	-34.6**	-34.9	-30.7
	GRAP_IFS	51.5**	81.8**	133.5*	186.6**	339.8	-22.5	-26.5	-27.8	-28.5	-26.9
	GRAP_KIM	73.7*	105.4	161.4	265.3	500.0	-21.5	-26.2	-28.0	-30.4	-28.5
	GRAP_UMM	71.2	110.1	172.5	251.6	391.1	-21.7	-26.2	-28.7	-28.7	-26.2
	PANG_IFS	54.0**	82.5*	135.1	208.1	318.0*	-24.3	-26.6	-27.7	-28.7	-28.3
	PANG_KIM	77.3**	110.7	193.7*	320.0**	569.1*	-22.9	-26.1	-28.2	-30.5	-30.0
	PANG_UMM	72.2	110.4	188.1	281.2	460	-24.0	-27.1	-29.0	-30.9	-29.0
SMS	FENW_IFS	54.7**	80.2**	130.3*	196.5*	314.7*	-	-	-	-	-
	FUXI_IFS	64.5	78.5*	132.2	195.3	337.3	-	-	-	-	-
MEAN±STD		64.3±12.2	97.4±20.5	159.4±35.9	246.0±58.2	416.7±91.5	-23.3±2.1	-27.1±2.3	-29.1±2.6	-30.5±2.8	-28.5±2.7

Table S2. Same as Table S1 but grouped by AI-WFM.

	DPE_ME					MWS_ME				
	Forecast lead times					Forecast lead times				
AI models	24	48	72	96	120	24	48	72	96	120
AIFS	63.3	105.1	170.7	283.1	479.1	-18.1**	-21.6**	-23.8**	-24.9*	-24.5
FENW	53.3**	77.6**	130.3**	203.0**	330.0**	-23.7	-27.4	-30	-33.5	-34.2
FGQG	56.6	87.3	123.9*	209.7	402.5	-25.2	-29.3	-31.6	-33.2	-31.1
FOUR	82.8**	130.0**	217.5**	329.8**	521.9*	-26.6**	-30.8**	-33.6**	-34.7**	-31.4
FUXI	59.1	75.9**	126.7**	188.6**	314.4*	-22.5	-25.8	-27.2	-26.9	-25.1
GRAP	61.4	95.7	150.4	225.0*	392.8	-21.9*	-26.4	-28.3	-29.2	-26.9
PANG	64.4	96.2	162.4	254.5	438.7	-23.2	-26.2	-27.9	-29.6	-27.9
MEAN ±STD	63.0 ±8.8	95.4± 17.1	154.6± 30.7	242.0± 46.7	411.3± 69.7	-23.0 ±2.5	-26.8 ±2.7	-28.9 ±2.9	-30.3 ±3.4	-28.7 ±3.3

Table S3. Same as Table S1 but grouped by IC.

	DPE_ME					WS_ME				
	Forecast lead times					Forecast lead times				
	24	48	72	96	120	24	48	72	96	120
GSM	63.4	104.5	170.7	274.9	464.9	-20.7**	-25.1*	-27.3	-28.3	-26.2
IFS	55.7**	82.3**	135.2**	205.5**	352.9**	-23.7	-27.1	-28.8	-30.0	-28.5
KIM	78.4**	113.4**	183.2**	293.9**	530.0**	-23.5	-27.5	-29.8	-32.0	-29.8
UMM	79.8**	124.2**	204.7**	302.6**	466.6	-24.1	-28.3	-30.8	-31.5	-28.6
MEAN ±STD	69.3±10.1	106.1 ±15.4	173.4 ±25.2	269.2 ±38.1	453.6 ±63.8	-23.0 ±1.3	-27.0 ±1.2	-29.2 ±1.3	-30.5 ±1.4	-28.3 ±1.3

Table S4. List of AI forecast model combinations that share identical AI-WFM architectures and IC sources across different contributing organizations.

	Organizations	AI models	Initial Condition
1	HKO, KMA	PANG	IFS
2	HKO, SMS	FENW	IFS
3	HKO, SMS	FUXI	IFS
4	HKO, KMA	GRAP	IFS

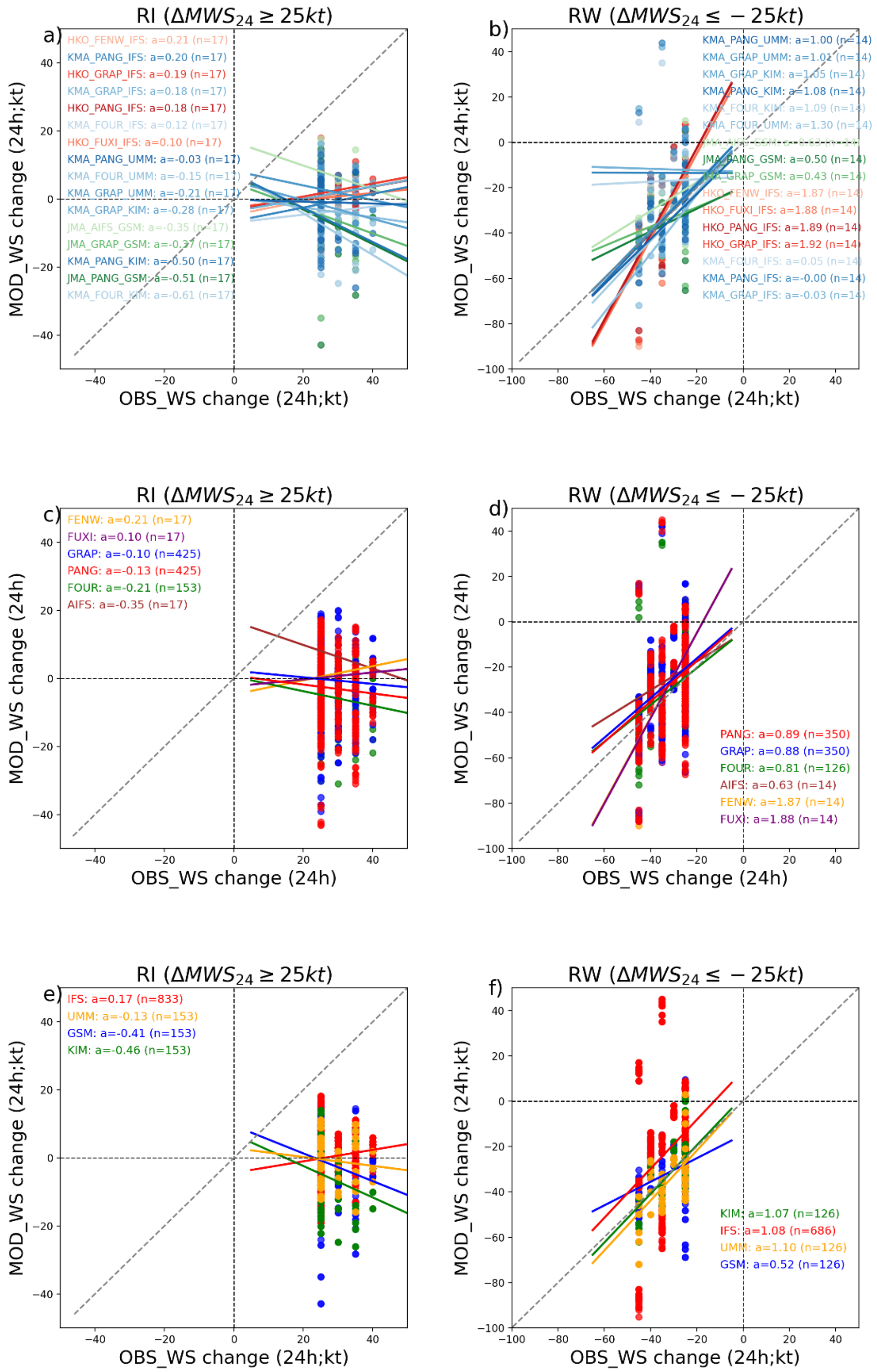


Figure S1. Scatterplots of observed versus forecast 24-h intensity change for (left) rapid intensification (RI) and (right) rapid weakening (RW), evaluated for (a, b) individual AI models, (c, d) AI-WFM groups, and (e, f) IC groups. Regression slopes (a) and sample sizes (n) are indicated for each model/group. Slopes are annotated in order of closeness to 1, indicating relative agreement between forecasted and observed changes.

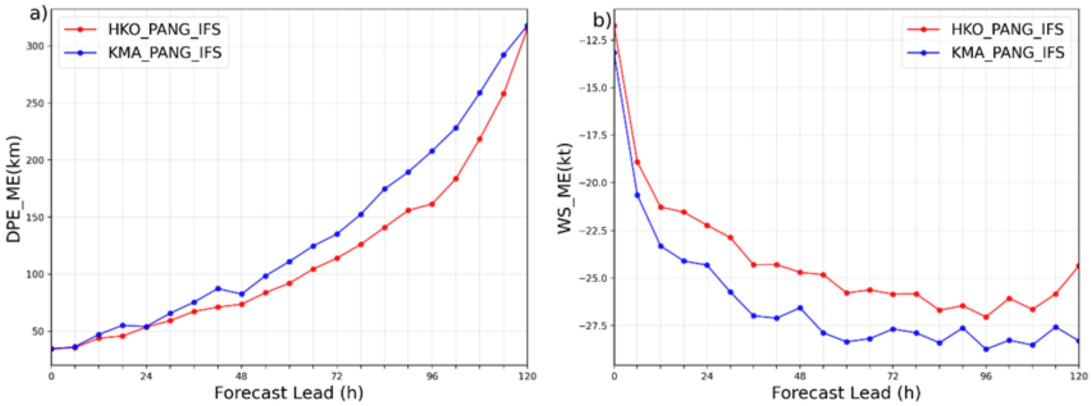


Figure S2. Comparison of forecast performance between HKO_PANG_IFS and KMA_PANG_IFS, which share the same AI system and IC configurations but exhibit substantially different track and intensity error characteristics. a) mean direct position error (DPE_ME, km); b) mean maximum wind speed bias (WS_ME, kt)