Supplementary Information: FuXi-TC

Shan Guo[2†], Lei Chen[1,2†], Yangyang Zhao[1,2], Yuetan Lin[1], Zeyi Niu[3,4], Xinyan Zhang[3], Ziyao Sun[3,4], Xiaohui Zhong[1,2*], Hao Li[1,2*]

[1]Artificial Intelligence Innovation and Incubation Institute, Fudan University, Shanghai, 200433, China.
[2]Shanghai Academy of Artificial Intelligence for Science, Shanghai, 200232, China.
[3]Department of Atmospheric and Oceanic Sciences, Institute of Atmospheric Sciences, Fudan University, Shanghai, 200433, China.
[4]Shanghai Typhoon Institute, China Meteorological Administration, Shanghai, 200030, China.

*Corresponding author(s). E-mail(s): x7zhong@gmail.com; lihao_lh@fudan.edu.cn;
Contributing authors: guoshan@sais.org.cn; cltpys@163.com; zhaoyangyang@sais.org.cn; linyuetan@sais.org.cn; niuzy@typhoon.org.cn; zhangxy@typhoon.org.cn; sunzy@typhoon.org.cn;
†These authors contributed equally to this work.

1 Tropical cyclone cases

Table 1 lists the tropical cyclone (TC) cases evaluated in both the Western North Pacific (WNP) and North Atlantic (NA) basins. For each storm, the storm identifier (SID) from International Best Track Archive for Climate Stewardship (IBTrACS) and the corresponding first and last forecast initialization times (UTC) are provided. All cases occurred in 2024. The inclusion of 17 WNP and 15 NA storms ensures coverage of diverse dynamical regimes and basin-dependent characteristics. By evaluating multiple initialization times for each TC, we systematically assess forecast performance across different lead times and storm development stages, thereby enabling a comprehensive validation of the proposed framework.

Table 1 List of TC names, TC SID and their first and last initialization times (in UTC) in WNP and NA Basins evaluated in this study. All dates refer to the year 2024.

Basin	TC Name	TC SID	First Init Time	Last Init Time	Init Times
WNP	GAEMI	2024201N12133	20 Jul, 1200	25 Jul, 1200	11
	PRAPIROON	2024201N14118	21 Jul, 1200	22 Jul, 1200	3
	MARIA	2024218N24141	08 Aug, 0000	12 Aug, 0000	9
	SON-TINH	2024224N27154	12 Aug, 0000	12 Aug, 0000	1
	AMPIL	2024225N22135	13 Aug, 0000	18 Aug, 0000	11
	WUKONG	2024225N24147	13 Aug, 1200	13 Aug, 1200	1
	JONGDARI	2024231N24126	19 Aug, 0000	19 Aug, 1200	2
	SHANSHAN	2024233N15148	22 Aug, 0000	29 Aug, 0000	15
	YAGI	2024245N12129	01 Sep, 1200	07 Sep, 1200	13
	LEEPI	2024246N22147	05 Sep, 1200	05 Sep, 1200	1
	BEBINCA	2024254N10148	10 Sep, 1200	16 Sep, 1200	13
	PULASAN	2024259N12145	15 Sep, 1200	20 Sep, 1200	11
	JEBI	2024270N15149	27 Sep, 1200	01 Oct, 1200	8
	KRATHON	2024271N22127	28 Sep, 0000	02 Oct, 1200	10
	BARIJAT	2024279N12148	09 Oct, 1200	10 Oct, 1200	3
	TRAMI	2024293N13141	22 Oct, 0000	27 Oct, 0000	11
	KONG-REY	2024298N13150	25 Oct, 0000	31 Oct, 1200	13
NA	BERYL	2024181N09320	29 Jun, 0000	08 Jul, 1200	20
	DEBBY	2024216N20284	04 Aug, 0000	09 Aug, 1200	11
	ERNESTO	2024225N14313	13 Aug, 0000	20 Aug, 0000	15
	FRANCINE	2024253N21266	10 Sep, 0000	12 Sep, 1200	6
	GORDON	2024256N16332	13 Sep, 1200	17 Sep, 0000	8
	HELENE	2024268N17278	25 Sep, 0000	28 Sep, 0000	7
	ISAAC	2024269N39302	26 Sep, 1200	30 Sep, 1200	9
	JOYCE	2024271N17319	27 Sep, 1200	29 Sep, 1200	5
	KIRK	2024274N14328	01 Oct, 0000	08 Oct, 0000	15
	LESLIE	2024275N10336	03 Oct, 0000	12 Oct, 0000	19
	MILTON	2024279N21265	06 Oct, 0000	11 Oct, 0000	11
	NADINE	2024293N17275	19 Oct, 1200	20 Oct, 0000	2
	OSCAR	2024293N21294	20 Oct, 0000	22 Oct, 0000	5
	PATTY	2024306N39319	02 Nov, 1200	04 Nov, 0000	3
	RAFAEL	2024309N13283	05 Nov, 0000	10 Nov, 0000	11

2 Comparison between WRF-° and WRF-°

We evaluated the model sensitivity to horizontal resolution by comparing simulations at 0.25° and 0.1° using identical physical configurations from April to October in 2023. Supplementary Figure 1 illustrates while track forecast accuracy remains similar between the two resolutions, the WRF-0.1° configuration demonstrates significantly superior performance in predicting TC intensity compared to WRF-0.25°.

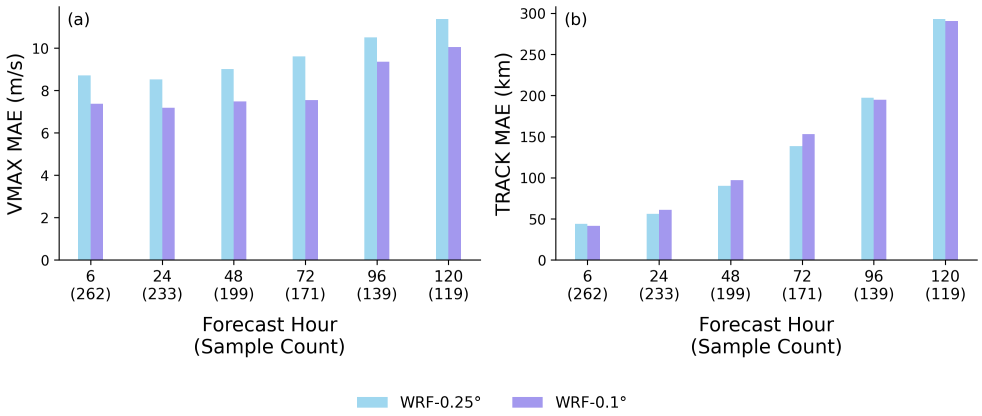

Fig. 1 Comparison of TC intensity and track forecasts between WRF simulations at different resolutions with IBTrACS dataset as ground truth. Mean absolute error (MAE) of maximum sustained wind and tracks with forecast lead time for WRF-0.25° (blue columns), WRF-0.1° (purple columns).

3 Spatial distribution of observed precipitation event counts

Figure 2 illustrates the spatial distribution of Automatic Weather Station (AWS)-based precipitation event statistics for rainfall exceeding 0.1 mm from July to September 2024 and their temporal variations across different forecast lead times. Given that the experiments were initialized at 00 and 12 UTC, this figure reveals that the volume of precipitation data received from AWS at 00 and 12 UTC is significantly larger than that available at 06 and 18 UTC.

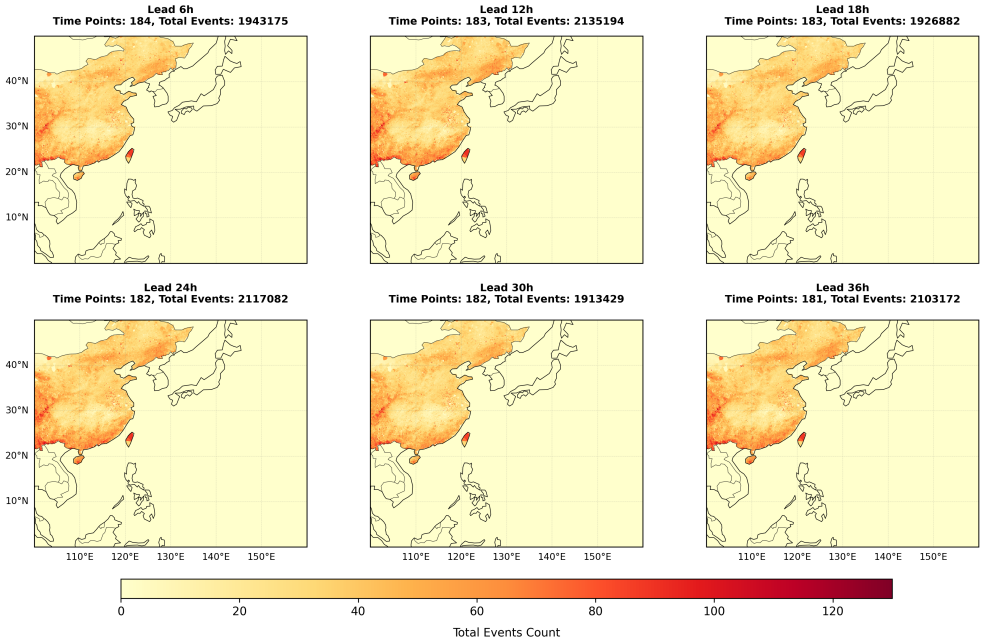

Fig. 2 Spatial distribution of observed precipitation event counts with a threshold of 0.1 mm by forecast lead time during July–September 2024.

4 Prediction of 2024 Typhoon Bebinca

Supplementary Figure 3 shows the 24-hour accumulated precipitation observation from AWS and forecasts initialized at 12 UTC on 14 September 2024. The comparison covers forecast lead times of 24-48, 48-72, 72-96 and 96-120 hours. These precipitation events encompass the rainfall over land induced by Typhoon Bebinca and Tropical Storm Pulasan. Notably, Bebinca was the strongest typhoon to make landfall in Shanghai since 1949, followed by Tropical Storm Pulasan, which made a subsequent landfall in Shanghai just four days later [1]. According to observations, Bebinca triggered heavy to torrential rainfall extending from the eastern coast to Central China, while Pulasan produced heavy rainfall with localized torrential downpours along the southeastern coast. Regarding models' performance, the FuXi model generally underestimates precipitation intensity. HRES-0.1° exhibits spatial displacement during the 48–72 h and 72–96 h intervals. However, FuXi-TC-0.1° inherits the precise spatial positioning of FuXi while demonstrating significant improvements in capturing precipitation intensity.

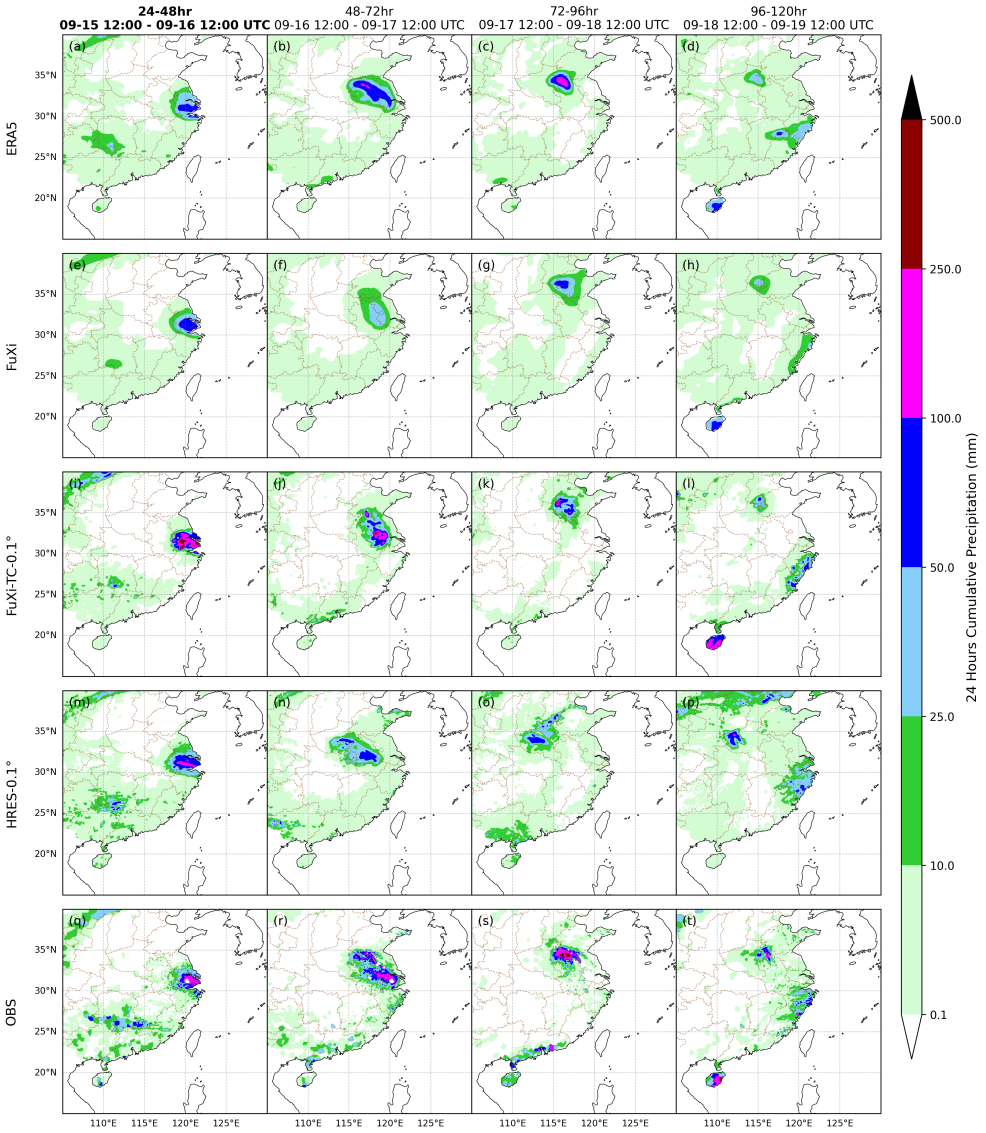

Fig. 3 Comparison of total precipitation (TP) for Typhoon Bebinca and Tropical Storm Pulasan forecasts. Comparison of snapshot examples of 24-hour accumulated TP among ERA5 (first row), FuXi (second row), FuXi-TC-0.1° (third row), HRES-0.1° (fourth row) and the AWS observations (fifth row) mainland China for 24-48 (first column), 48-72 (second column), 72-96 (third column) and 96-120 (fourth column) hours forecasts with the initial forecasting time at 12:00 UTC 14 September 2024.

Supplementary Figure 4 presents a comparison of forecast tracks for Typhoon Bebinca among FuXi-ENS, FuXi-TC-ENS, and ECMWF-ENS across different initialization times. Notably, the differences between the ensemble mean tracks of FuXi-TC-ENS and FuXi-ENS are negligible. This is attributed to the fact that the post-processing framework primarily adjusts typhoon intensity and exerts minimal influence on the track. Regarding TC track forecasts, ECMWF-ENS exhibits slightly inferior performance in this case study compared to both FuXi-ENS and FuXi-TC-ENS, particularly for forecasts initialized 5 to 6 days in advance.

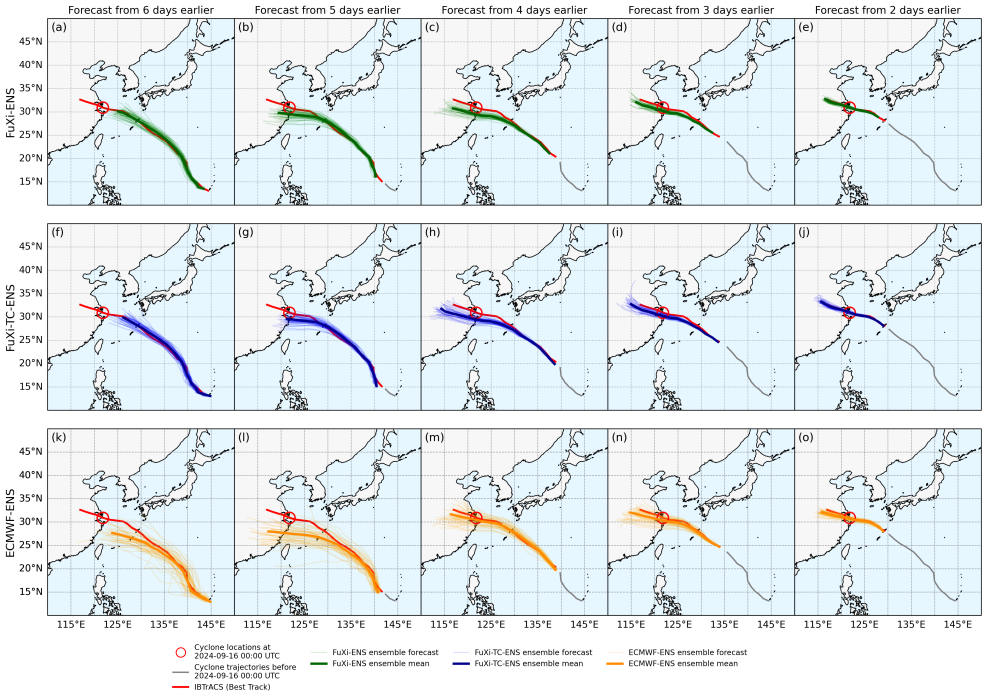

Fig. 4 Tracking results for Typhoon Bebinca from ERA5 (red line) and ensemble forecasts generated by FuXi-ENS (first row, green line), FuXi-TC-ENS (second row, blue line) and ECMWF-ENS (third row, orange line). The columns show forecasts initialized at 12:00 UTC on 10, 11, 12, 13, and 14 September 2024 (from left to right), corresponding to 6, 5, 4, 3, and 2 days before landfall, respectively. The gray line represents the observed track prior to each forecast initialization time, and the red circle indicates the landfall location.

5 Overall precipitation evaluation in WNP

Supplementary Figure 5 and 6 present the time series of Probability of Detection (POD) and False Alarm Ratio (FAR) for 6-hour accumulated TP across varying thresholds from 0.1 mm to 40 mm. POD reflects the model's sensitivity to observed precipitation events, while FAR indicates the propensity for false alarms. In terms of model performance, FuXi performs well at the lowest threshold. However, at higher thresholds, both its POD and FAR decrease significantly, a trend attributed to the model's general tendency to underestimate precipitation intensity. Notably, in the evaluation for TP at 40 mm threshold, FuXi's FAR exhibits large fluctuations, likely due to the insufficient sample size of extreme precipitation events predicted by FuXi. Comparatively, HRES-0.1° demonstrates a relatively stable trade-off between POD and FAR across thresholds. Both WRF-0.1° and FuXi-TC-0.1° show superiority at higher precipitation thresholds. WRF-0.1° exhibits a higher FAR at thresholds above 0.1 mm, possibly due to a systematic wet bias in the numerical model. During the initial forecast period, both POD and FAR are relatively low for WRF-0.1° , reflecting the spin-up process during which precipitation is not fully developed.

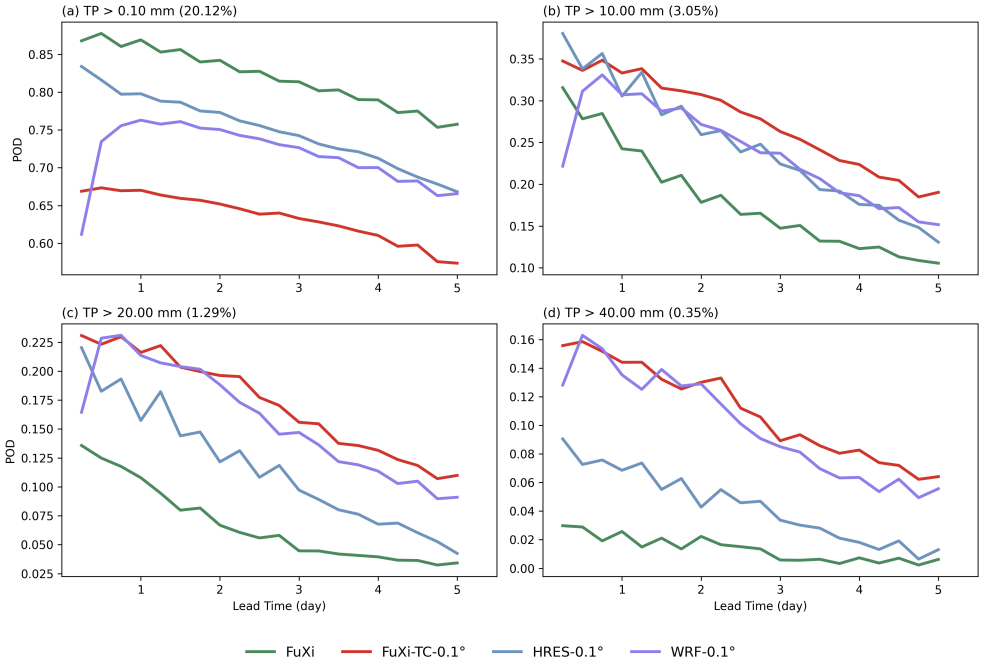

Fig. 5 Comparison of POD for TP. Comparison for the FuXi(green lines), FuXi-TC-0.1°(red lines), HRES-0.1°(blue lines) and WRF-0.1°(purple lines) in predicting 6-hour precipitation for 0.1, 10, 20, 40 mm thresholds in (a-d) respectively. The ratios of extreme cases relative to the entire test set are indicated in parentheses. All the forecast data are evaluated against the Automatic Weather Station observations in period 1 July to 30 September 2024.

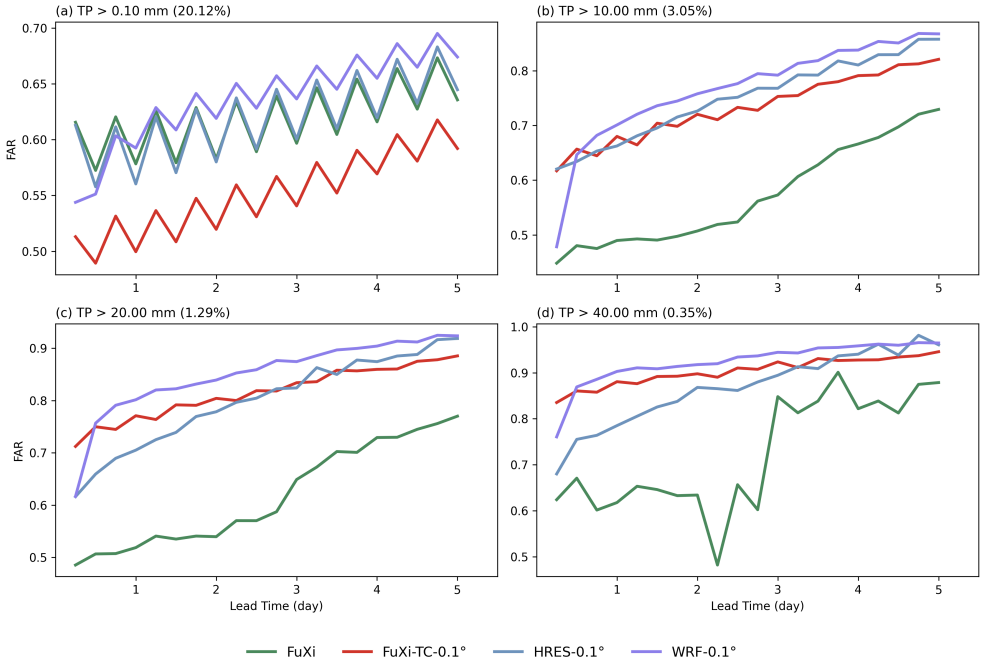

Fig. 6 Comparison of FAR for TP. Comparison for the FuXi (green lines), FuXi-TC-0.1° (red lines), HRES-0.1° (blue lines) and WRF-0.1° (purple lines) in predicting 6-hour precipitation for 0.1, 10, 20, 40 mm thresholds in (a-d) respectively. The ratios of extreme cases relative to the entire test set are indicated in parentheses. All the forecast data are evaluated against the Automatic Weather Station observations in period 1 July to 30 September 2024.

6 Energy spectra

We employed kinetic energy spectra to evaluate the effective resolution and smoothness of the model outputs. By decomposing the kinetic energy into spectral components, we assessed the model's ability to preserve variance across different scales. Generally, spectral energy exhibits a power-law decay as wavelengths decrease, reflecting the energy cascade from organized large-scale circulation to smaller-scale turbulent processes. Supplementary Figure 7 presents energy spectra within the WNP domain for 10-meter wind speed (WS10M) and mean sea level pressure (MSL) across 24, 72, 120 hours forecast lead times. The 0.1°-resolution models exhibit distinct spectral characteristics compared to their 0.25° counterparts. FuXi achieves remarkable alignment with European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5) across all scales and surface variables. Notably, FuXi-TC-0.1° significantly outperforms the FuXi. At smaller wavelengths, FuXi-TC-0.1° initially aligns with WRF-0.1° but shows slightly reduced energy with increasing wavenumber, suggesting smoother forecasts and a loss of fine-scale details for WS10M. However, for MSL, FuXi-TC-0.1° indicates superior consistency with WRF-0.1°, suggesting that the model is more capable of capturing small-scale features and details for this variable.

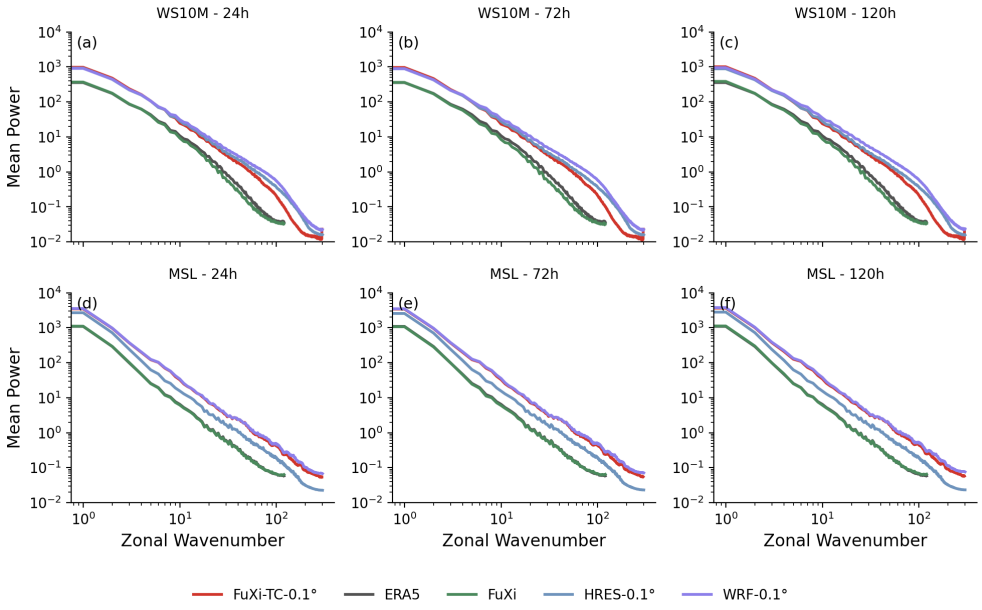

Fig. 7 Energy spectra of FuXi-TC-0.1°(red lines), ERA5(black lines), FuXi(green lines), HRES-0.1°(blue lines) and WRF-0.1°(purple lines). The figure includes WS10M (first row) and MSL (second row). The spectra are shown at forecast lead times of 24 hours (first column), 72 hours (second column) and 120 hours (third column), using the dataset from July to October 2024.

7 Overall MSL evaluation in WNP and NA

Supplementary Figure 8 presents the overall MAE of the MSL for TCs in both the WNP and NA basins. In the WNP, FuXi-TC-0.1° demonstrates a significant improvement in MSL accuracy over the original FuXi model across all lead times. For lead times within the first 48 hours, its performance is comparable to that of WRF-0.1°. However, since WRF-0.1° exhibits a higher MAE than HRES-0.1°, HRES-0.1° remains the superior performer overall. The improvement in MSL accuracy relative to FuXi remains robust in different regions, which demonstrates that the FuXi-TC framework can be directly applied to the NA basin without retraining using WRF-0.1° data.

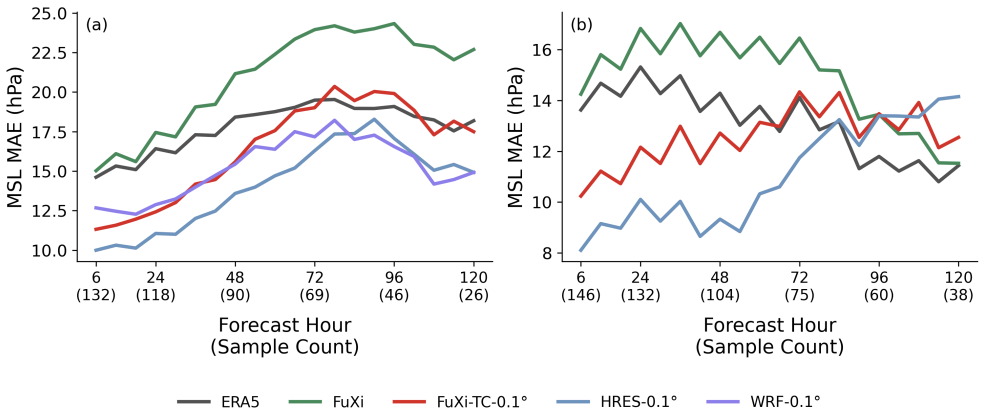

Fig. 8 Comparison of the average MAE for MSL in WNP (a) and NA (b) with forecast lead time for ERA5 (black lines), FuXi (green lines), FuXi-TC-0.1° (red lines), HRES-0.1° (blue lines) and WRF-0.1° (purple lines). All the forecast data are evaluated against the IBTrACS dataset. The evaluation covers forecasts for 17 TCs in WNP from July to October 2024, and 15 TCs in NA from June to December.

8 Wind speed prediction comparison of tropical cyclones cases in NA

We present the forecast evolution characteristics of WS10M for Hurricanes Beryl (initial forecasting time at 12:00 UTC, 2 July 2024) and Helene (initial forecasting time at 00:00 UTC, 25 September 2024) in Figures 9 and 10. The three rows of the figures sequentially compare the forecast snapshots of ERA5, FuXi, and FuXi-TC-0.1°, while the four columns correspond to the 12-, 24-, 36-, and 48-hour forecast results, respectively. It can be observed that FuXi-TC-0.1° significantly increase the predicted values of WS10M while maintaining the consistency of the cyclone center structure with FuXi.

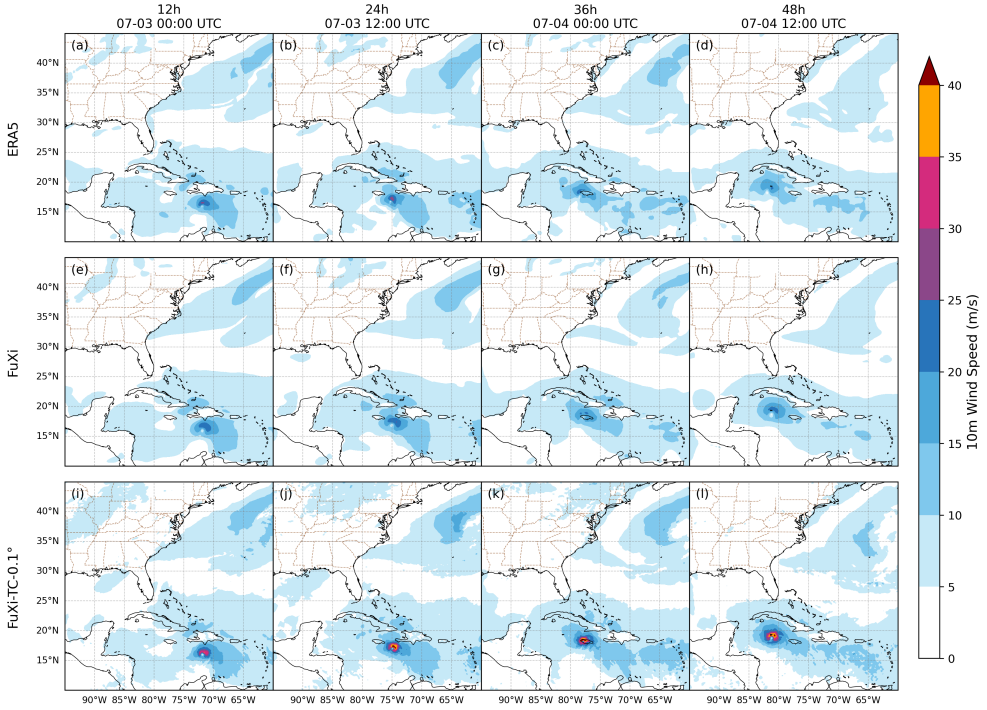

Fig. 9 Comparison of WS10M for Hurricane Beryl forecast. Comparison of snapshot examples of WS10M among ERA5 (first row), FuXi (second row) and FuXi-TC-0.1° (third row) for 12 (first column), 24 (second column), 36 (third column) and 48 (fourth row) hours forecasts with the initial forecasting time at 12:00 UTC, 2 July, 2024.

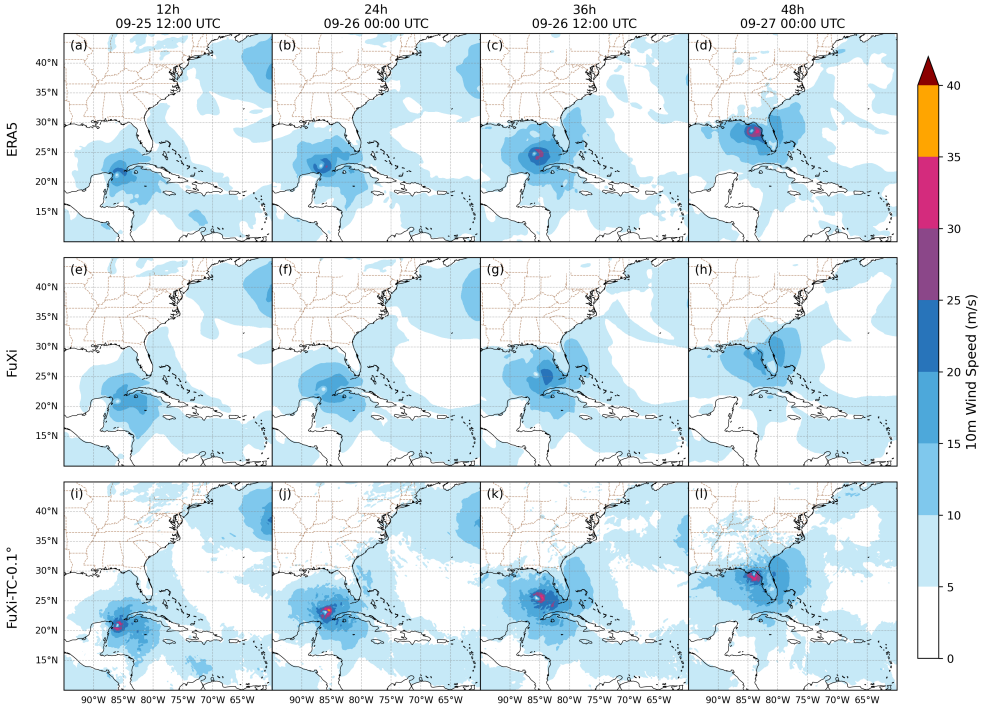

Fig. 10 Comparison of WS10M for Hurricane Helene forecast. Comparison of snapshot examples of WS10M among ERA5 (first row), FuXi (second row) and FuXi-TC-0.1° (third row) for 12 (first column), 24 (second column), 36 (third column) and 48 (fourth row) hours forecasts with the initial forecasting time at 00:00 UTC, 25 September, 2024.

9 Domain setup and TC track forecasts in the WNP and NA Basins

Supplementary Figure 11 illustrates the domain configurations used for the WNP and NA basins. To accommodate the broader longitudinal span of the NA basin while maintaining consistent input dimensions, the NA domain is constructed by horizontally concatenating two sub-domains. Each of these NA sub-domains is identical in size to the WNP domain. Specifically, the western NA sub-domain spans 99.2°W–40.1°W and 5.5°N–55°N while the eastern NA sub-domain spans 59.8°W–0.1°W and 5.5°N–55°N. In the overlapping region between the two NA sub-domains, the final meteorological fields are obtained by calculating the average of the overlapping grid points.

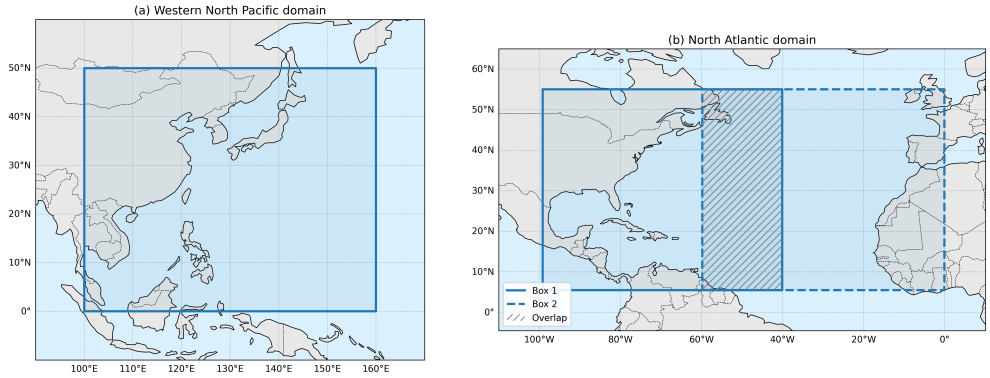

Fig. 11 Domains used in this study. (a) The Western North Pacific (WNP) domain, (b) The North Atlantic (NA) domain. The NA domain is constructed by horizontally concatenating two overlapping sub-domains with identical size (Box 1, solid; Box 2, dashed), with the hatched area indicating the overlap region.

Within these defined domains, supplementary Figure 12 presents the track forecast comparisons for four selected tropical cyclones cases. The first row displays Typhoons Bebinca and Gaemi over the WNP, while the second row shows Hurricanes Beryl and Helene over the NA. The results indicate that the track predictions of FuXi-TC-0.1° (red line) are nearly identical to those of the baseline FuXi model (green line). For the late-stage forecasts of Hurricane Beryl, the deviation of HRES-0.1° (light blue line) from the International Best Track (black line) gradually increases with the extension of the forecast lead time. Regarding Hurricanes Helene, the track forecasts generated by the FuXi-series models exhibit slightly larger deviations in the predicted landfall location.

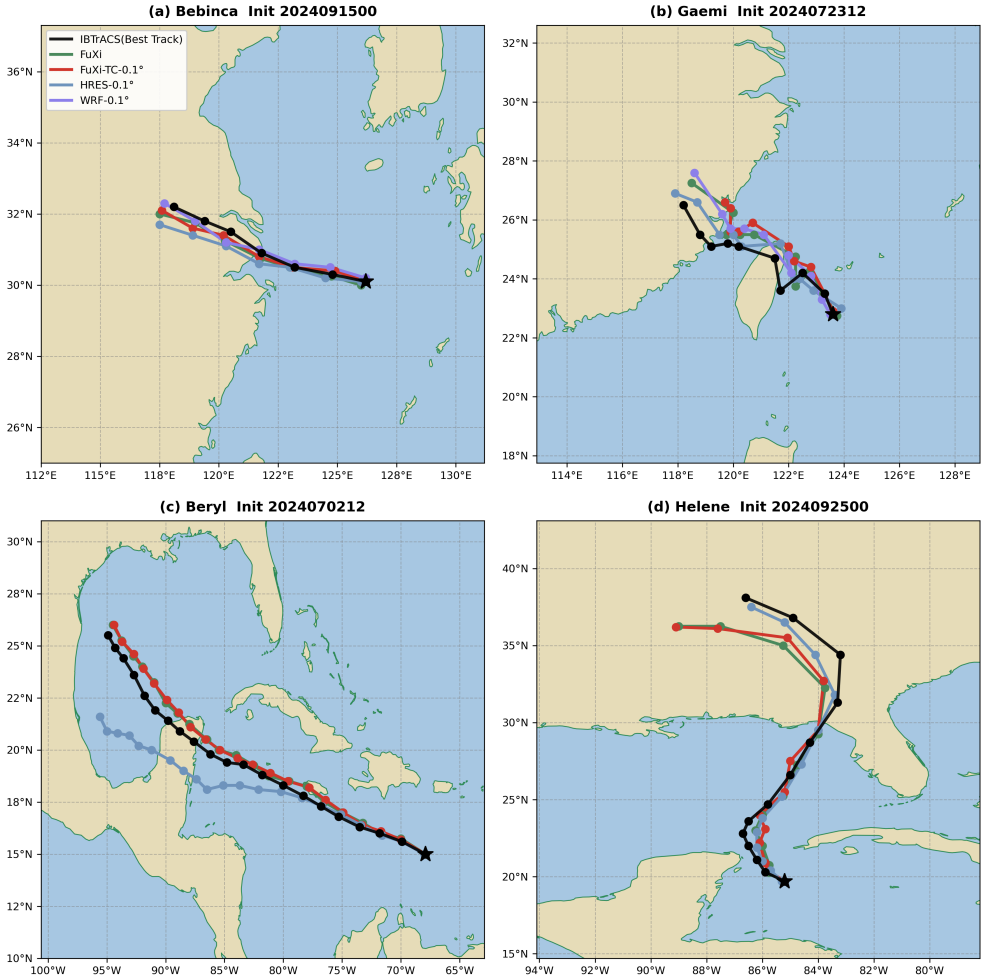

Fig. 12 Track comparison of Typhoons (a) Bebinca, (b) Gaemi (c) Beryl and (d) Helene. Tracking results from FuXi (green line), FuXi-TC-0.1° (red line), HRES-0.1° (light blue line), and the observed Best Track (black line) are presented as lines with markers.

References

[1] Wu, L., Yu, R., Xiang, C., Yu, H., Feng, Y., Zhou, X.: Extreme Impacts of Four Landfalling Tropical Cyclones in China during the 2024 Peak Season. Springer (2025)