PhysTC: A Physics-Enhanced Dataset and Architecture for High-Precision Tropical Cyclone Forecasting

Zhaoran Feng[1,2] · Xuanhong Chen[3*] · Zeng-Bing Chen[1*] · Shengjun Wu[1] · Bin He[4*] · Kairui Feng[2,4*]

[1] Nanjing University [2] Shanghai Innovation Institute [3] Shanghai Jiao Tong University

[4] Shanghai Research Institute for Intelligent Autonomous Systems & State Key Laboratory of Autonomous Intelligent Unmanned Systems, Tongji University

Feel free to contact me

Zhaoran Feng

602025220012@smail.nju.edu.cn

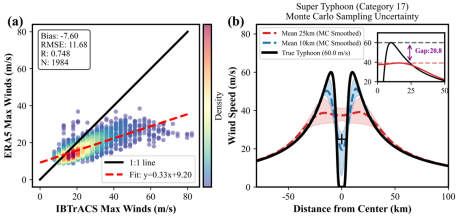
* Corresponding authors

NEURAL INFORMATION PROCESSING SYSTEMS

NeurIPS 2026

ED Track

Why a New Benchmark?



Coarse inputs miss extreme intensity

Grid averaging suppresses peak winds. A 25 km filter reduces a 60 m/s vortex peak by more than 20 m/s.

PhysTC pairs coarse environmental fields with best-track targets.

Data Construction

Time alignment

Match ERA5 and IBTrACS on a 6-hour grid at 00/06/12/18 UTC. Extract storm-centered environmental fields.

Label harmonization

Standardize intensity labels to 10-minute sustained winds for consistent cross-basin evaluation.

Quality control

Log missing timestamps and exclude incomplete ERA5 or best-track samples without interpolation.

Split integrity

Group training/validation storms in 1950–2016. Reserve 2017–2023 for held-out testing.

Dataset & Benchmark

Global tropical cyclones, 1950–2023, aligned at 6-hour intervals

ERA5 environmental fields paired with IBTrACS best-track supervision

Dataset property	Dig. Typh.	IBTrACS	WB2	SHIPS	TCND	PhysTC
Environmental grids	No	No	Yes	Yes	Yes	**Yes**
Physical cues	No	No	No	Yes	No	**Yes**
Scale bridging	No	No	No	No	No	**Yes**
Global scope	No	Yes	Yes	No	Yes	**Yes**
Clean split	No	No	Yes	No	Yes	**Yes**
TC protocol	No	No	No	Yes	Yes	**Yes**

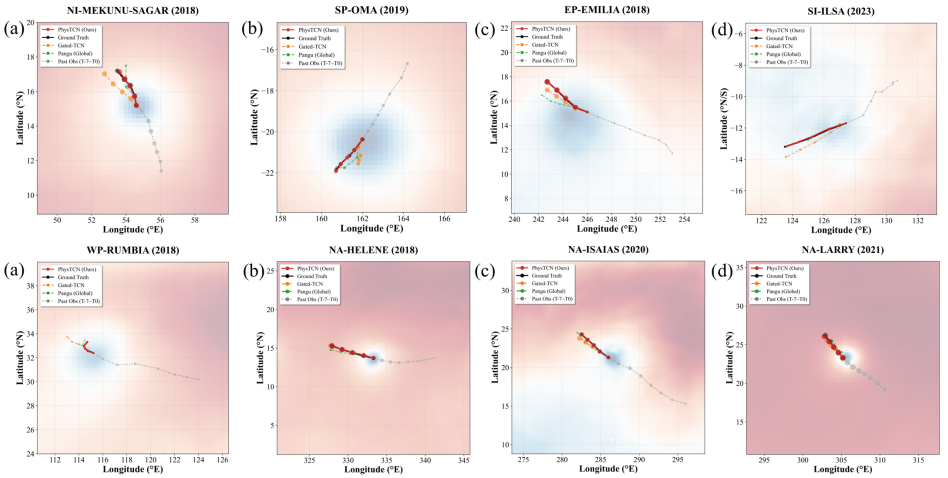
Selected dataset properties reported in paper Table 2.

Eight observed frames predict track, wind and pressure at 6, 12, 18 and 24 h.

Evaluation spans six basins, with tail-intensity and rapid-intensification tests.

Cross-Basin Forecasting

Representative track forecasts on held-out storms from six ocean basins

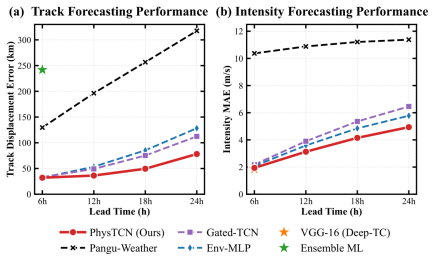


Quantitative Results

Global forecast errors, averaged over 6–24 h

Model	Track ADE (km)	Pressure MAE (hPa)	Wind MAE (m/s)
Pangu-Weather	225.06	11.64	10.96
HWRF/HWFI	140.56	—	5.12
Env-MLP	74.80	5.78	4.07
Gated-TCN	67.53	6.49	4.47
DeepTC (VGG-16)†	—	5.84	1.86
Ensemble-ML†	241.82	—	—
PhysTCN (Ours)	**49.57 ± 1.31**	**5.19 ± 0.08**	**3.57 ± 0.03**

PhysTCN table results: mean ± SD over 20 seeds.
† Additional baselines as reported in paper Table 3.



Contact

Zhaoran Feng

602025220012@smail.nju.edu.cn

Supported by

This work was supported by the **Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China** (Grant No. JYB2025XDXM908), the **National Natural Science Foundation of China** (Grant Nos. T2600685 and 92565111), and the **Xiaomi Foundation**. We also gratefully acknowledge support from the **Shanghai Innovation Institute**.