

<https://doi.org/10.1038/s43247-025-02347-5>

An artificial intelligence-based limited area model for forecasting of surface meteorological variables



Pengbo Xu¹, Xiaogu Zheng^{2,3}, Tianyan Gao⁴, Yu Wang¹, Junping Yin^{2,5,6}✉, Juan Zhang^{2,7},
Xuanze Zhang⁸, San Luo⁹, Zhonglei Wang¹⁰, Zhimin Zhang¹¹, Xiaoguang Hu¹¹ & Xiaoxu Chen¹¹

Artificial intelligence-based models for global weather forecasting have advanced rapidly, but research on high-resolution data-driven limited area models remains scarce. Here we introduce a limited area artificial intelligence-based weather forecasting model with 3 km and 1 h resolutions, utilizing parallel global-local structures to capture multiscale meteorological features. The model is trained on high-resolution regional analysis data, and utilizes the global artificial intelligence-based model forecasts for the lateral boundary condition during prediction, and operates much faster than the dynamical forecast model. In two selected limited areas, it outperforms dynamical forecast models in surface wind speed forecasting but underperforms in surface temperature and pressure. Skills in surface temperature and pressure can be further improved comparable to the dynamical forecast model by providing better lateral boundary conditions. Issues related to lateral boundary conditions, such as selecting width of lateral boundary regions and combining finer and coarser resolution predictions in the regions, are also studied.

Accurate weather forecasting with higher resolution plays a crucial role in modern society, especially for surface meteorological variables¹. Research in numerical weather prediction (NWP) has advanced rapidly over the past decades^{2,3}. Traditional NWP involves numerically solving the governing equations of fluid dynamics^{4,5}, which is often computationally expensive and time-consuming^{6–10}.

Recently, several artificial intelligence-based (AI-based) global weather forecasting models have been developed^{11–20}, such as FourCastNet²¹, Pangu-Weather²², GraphCast²³, FengWu²⁴, FuXi²⁵, ClimaX²⁶, Aurora²⁷, and AIFS²⁸. Most of these models are trained using global ERA5 reanalysis data²⁹ with a spatial resolution of $0.25^\circ \times 0.25^\circ$, matching the ECMWF Integrated Forecasting System³⁰. AI-based models consume much fewer computational resources and run much faster than NWP. Since ERA5 reanalysis data is generated by assimilating observations into NWP forecasts, it can correct some model errors related to NWP, such as those associated with the parameterization of physical processes³¹. Therefore, AI-based weather

forecasting models trained on ERA5 reanalysis data could potentially be more skillful than traditional NWP.

For weather forecasts with finer spatial resolution (such as 3 km) in domains ranging from several countries to a county or city, limited area models (LAMs) are developed. However, most work on LAMs is related to NWP, and studies on AI-based LAMs remain limited. The MetNet series^{32,33} is an attempt in this area, pre-trained using finer spatial resolution analysis data and focusing on precipitation, radar, satellite, and ground site observations. Building on the architecture of GraphCast, a graph-based neural model was developed³⁴ to generate forecasts for the Nordic region by fully replacing the lateral boundary region with NWP data at the corresponding time.

Recently, AI-based models with high resolution in a region have been established by training on both fine resolution regional analysis data and coarse resolution global reanalysis data. For example, StormCast³⁵ applied generative diffusion modeling trained with 3 km resolution analysis data

¹School of Mathematical Sciences, Key Laboratory of MEA (Ministry of Education), Shanghai Key Laboratory of PMMP, East China Normal University, Shanghai, PR China. ²Shanghai Zhangjiang Institute of Mathematics, Shanghai, PR China. ³International Global Change Institute, Hamilton, New Zealand. ⁴School of Mathematics and Statistics, Northeast Normal University, Changchun, PR China. ⁵Institute of Applied Physics and Computational Mathematics, Beijing, PR China. ⁶National Key Laboratory of Computational Physics, Beijing, PR China. ⁷Institute of Artificial Intelligence, Beihang University, Beijing, PR China. ⁸Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, PR China. ⁹China Meteorological Administration Basin Heavy Rainfall Key Laboratory/Hubei Key Laboratory for Heavy Rain Monitoring and Warning Research, Institute of Heavy Rain, China Meteorological Administration, Wuhan, PR China. ¹⁰Wang Yanan Institute for Studies in Economics and School of Economics, Xiamen University, Xiamen, Fujian, PR China. ¹¹Baidu Inc., Beijing, PR China. ✉e-mail: yinjp829829@126.com

from the High-Resolution Rapid Refresh (HRRR) and 0.25° resolution ERA5 data for high-resolution forecasts in Central America. A graph-based neural model with a global stretched grid was developed³⁶, and trained using 2.5 km resolution resolution operational analyses from the MetCoOp Ensemble Prediction System data³⁷ and ERA5 data for high-resolution forecasts in the Nordics. Recently, OMG-HD³⁸ used observation data from the past 6 hours from meteorological stations and satellites to train an AI-based LAM for forecasting surface meteorological variables on a resolution of $0.05^\circ \times 0.05^\circ$. However, due to the lack of an imposed lateral boundary condition (LBC), only 12-h effective forecasts are available.

In this study, an AI-based LAM, called YingLong, is developed for forecasting surface meteorological variables, including surface temperature, pressure, humidity, and wind speed. These variables are highly sensitive to human life, alongside precipitation. Unlike previous work, YingLong is trained solely using 3 km resolution HRRR regional analysis data. Although global coarse resolution data are not used for training, they are required to supply the LBC. This study investigates YingLong with LBC imposed by Pangu-weather for forecasting surface meteorological variables.

YingLong is also applied to investigate issues related to LBCs of AI-based LAMs. As is well known, the domain of a LAM is divided into an inner area and a surrounding lateral boundary region. The meteorological conditions on the lateral boundary region are provided by coarse-resolution models, while the inner area is the focus of research and application. The coarse-resolution atmospheric information on the lateral boundary region is transmitted to the inner area through the forecast of the high-resolution LAM. Numerical LAMs require extensive processing of coarse resolution LBCs to accommodate differences in the model structure and parameterization schemes, ensuring that certain physical constraints are met. However, this may not be necessary for AI-based LAMs. As statistical models, AI-based LAMs have fewer physical restrictions on the model structure and initial values compared to numerical models. Furthermore, effectively transferring meteorological information from the lateral boundary region to the inner area is also a key objective of this paper.

Results

This section is dedicated to the construction of YingLong and its predictability with LBC imposed by Pangu-weather for the two selected limited areas in North America.

Construction of YingLong

YingLong is an AI-based limited area weather forecasting model with a 3 km spatial resolution and 24 weather variables (see Table 1). It is trained on hourly HRRR analysis data (denoted as HRRR.A), created by assimilating observations into WRF-ARW (Advanced Research WRF³⁹) in the period 2015–2021, and validated and tested on HRRR.A in 2023 and 2022, respectively (see Methods “Datasets” for more details). Since the WRF-ARW forecasts are documented in HRRR dataset and used in data

Table 1 | The weather variables of YingLong include a total of 24 variables, consisting of 20 upper-air variables from 4 pressure levels (50 hPa, 500 hPa, 850 hPa, and 1000 hPa, following FourCastNet) and four surface variables

Type	Full name	Abbreviation
Upper air variables	Geopotential height	Z
	Specific humidity	S
	Temperature	T
	U component of wind speed	U
	V component of wind speed	V
Surface variables	Mean sea level pressure	MSLP
	2 m temperature	T2M
	U component of 10 m wind speed	U10
	V component of 10 m wind speed	V10

assimilation for HRRR.A, it is denoted as HRRR.F for convenience. To test YingLong’s sensitivity to topography, YingLong is constructed and evaluated on two domains: the east domain (ED), which is relatively flat, and the west domain (WD), which includes considerable mountain ranges (see Fig. 1). Both ED and WD are divided into an inner area (the area within the red dashed line) and a lateral boundary region (the region outside the red dashed line).

The architecture of YingLong includes an embedding layer, a spatial mixing layer (comprising the Local branch and Global branch), and a linear decoder (see Fig. 2 and Methods “Architecture” for details). YingLong is trained using the procedure documented in Methods “Training”, with a smoothing boundary condition imposed (see Methods “Lateral boundary condition”). The procedure for YingLong’s rolling forecasting is documented in Methods “Rolling forecast” and briefly illustrated in Fig. 3. Two indicators, including root mean squared error (RMSE, Supplementary Eq. (S1) in Supplementary Note 1) and anomaly correlation coefficient (ACC, Supplementary Eq. (S2) in Supplementary Note 1), are used to evaluate the models. Specifically, a smaller RMSE or a larger ACC indicates a better performing model.

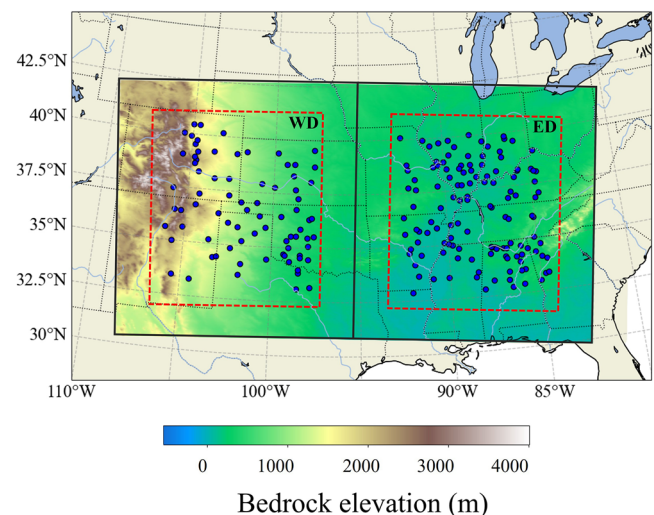


Fig. 1 | The topography of ED and WD. The dots represent observation sites. In total, there are 205 sites, 78 in WD and 127 in ED.

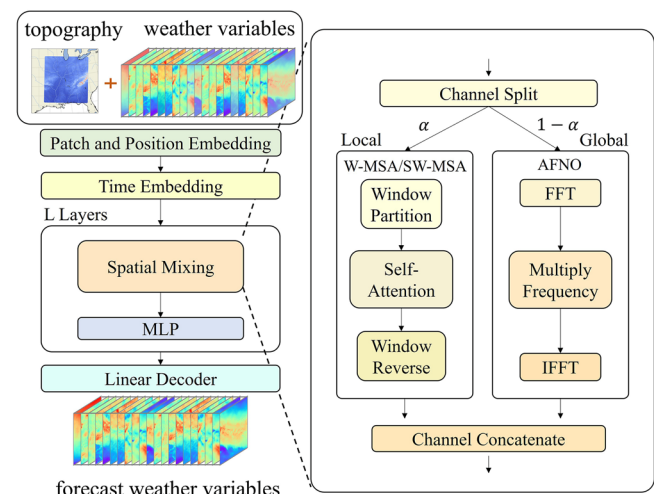


Fig. 2 | The architecture of YingLong. The architecture of YingLong includes an embedding layer, a spatial mixing layer composed of the Local branch and Global branch, and a linear decoder. FFT and IFFT represent the fast Fourier transform and inverse fast Fourier transform, respectively. See Methods “Architecture” for details.

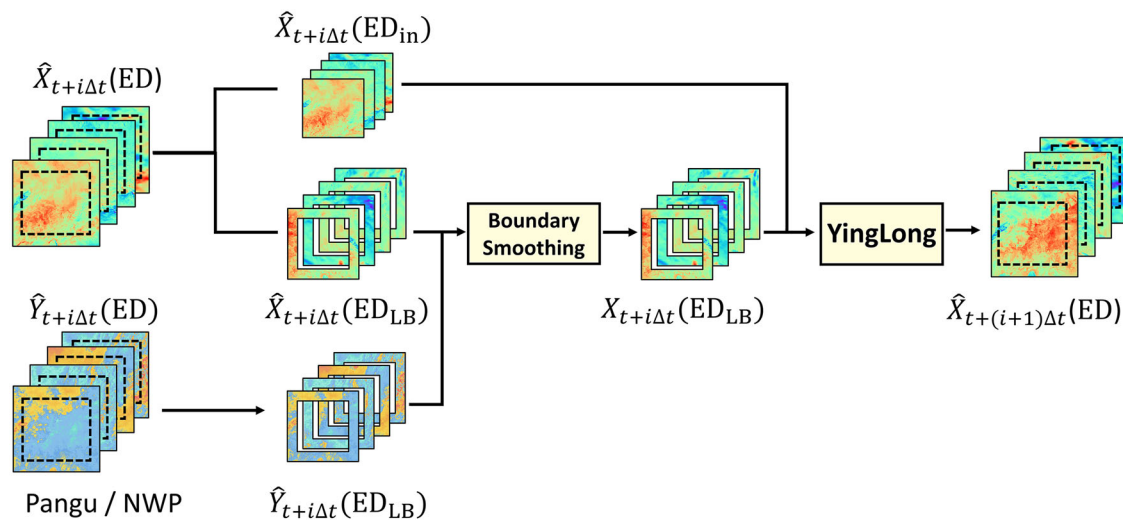


Fig. 3 | The framework of YingLong model. Rolling forecasts of YingLong on the region ED through the process in Methods “Rolling forecast”. The prediction of YingLong on the ED $\hat{X}_{t+i\Delta t}(\text{ED})$ is divided into $\hat{X}_{t+i\Delta t}(\text{ED}_{\text{in}})$ on the inner area and $\hat{X}_{t+i\Delta t}(\text{ED}_{\text{LB}})$ on the lateral boundary region. Then boundary smoothing (Eq. (3)) is

applied to $\hat{X}_{t+i\Delta t}(\text{ED}_{\text{LB}})$ and $\hat{Y}_{t+i\Delta t}(\text{ED}_{\text{LB}})$ (the forecast of Pangu on ED_{LB}) to form $X_{t+i\Delta t}(\text{ED}_{\text{LB}})$. Finally, $\hat{X}_{t+i\Delta t}(\text{ED}_{\text{in}})$ and $X_{t+i\Delta t}(\text{ED}_{\text{LB}})$ are merged for the initial value of YingLong prediction $\hat{X}_{t+(i+1)\Delta t}(\text{ED})$.

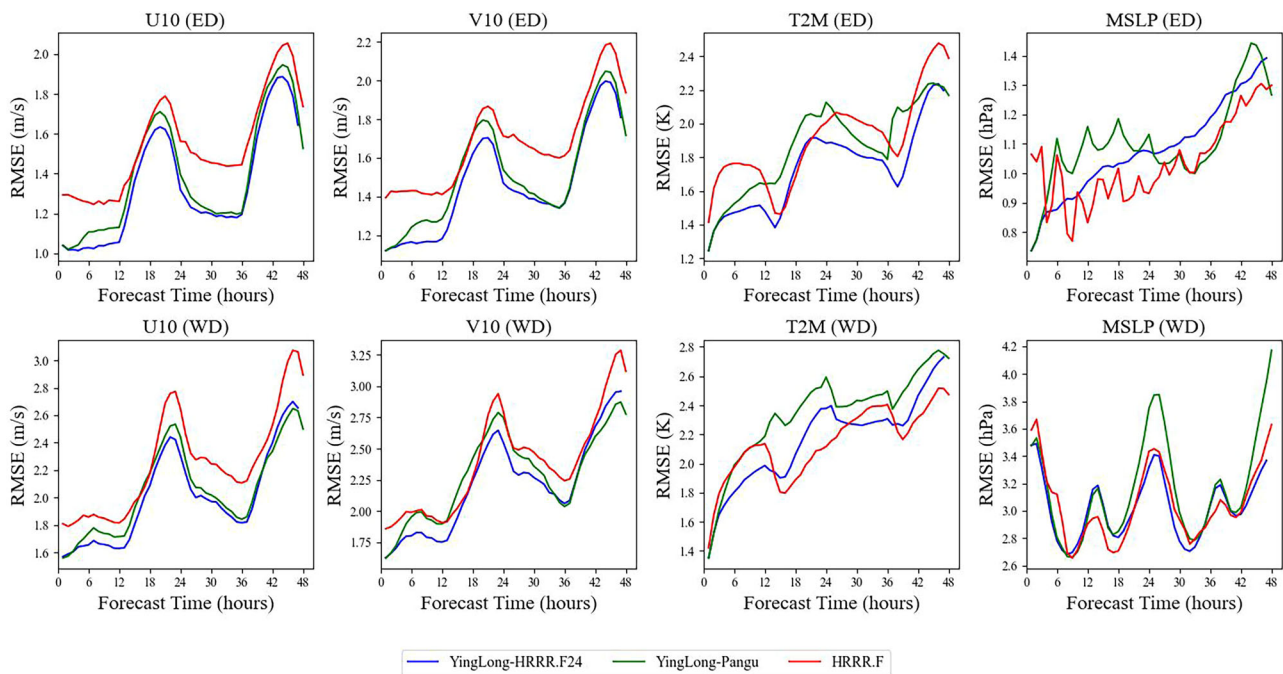


Fig. 4 | Forecast evaluations YingLong models. RMSEs with in-situ observations as benchmark of HRRR.F (red lines), YingLong-HRRR.F24 (blue lines), and YingLong-Pangu (green lines) in ED and WD.

Forecasting skills of YingLong-Pangu

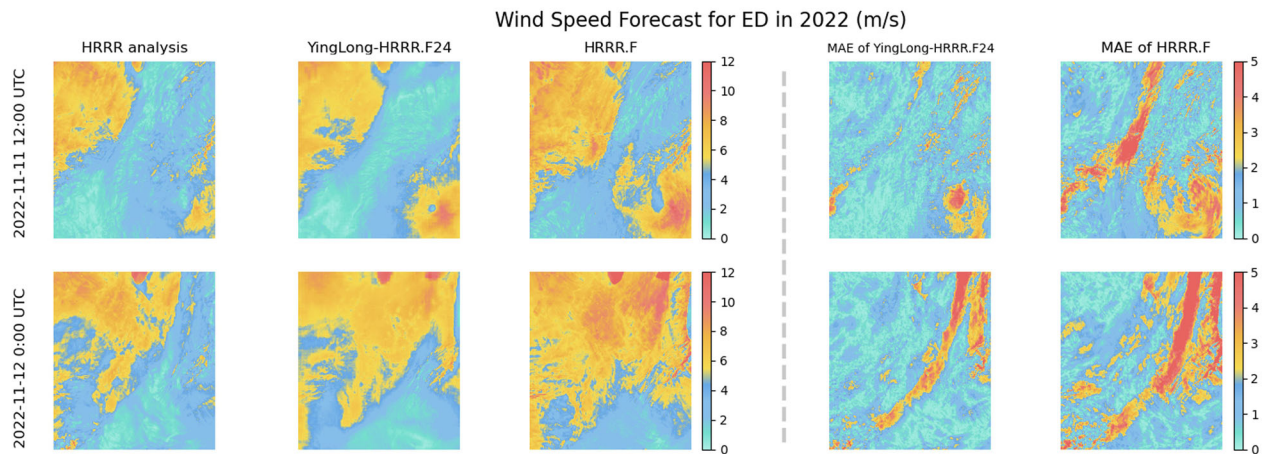
To construct a fully AI-based LAM with fine spatial resolution, the Pangu-weather medium-term forecasts are applied as the LBC of YingLong (denoted as YingLong-Pangu). To evaluate its forecasting skills more objectively, the in-situ observation data⁴⁰ is used as the benchmark to calculate the RMSEs of the forecasts. In this study, the forecast for a station observation is defined as the model forecast at the 3 km grid point containing the observation site.

Figure 4 shows the RMSEs of 48-hour rolling forecasts for YingLong-Pangu and HRRR.F, starting at 00:00 UTC, for surface variables (U10, V10, T2M, MSLP) at the stations in the inner areas of ED and WD in 2022. The results indicate that the RMSEs of YingLong-Pangu (green lines) are generally smaller than those of HRRR.F (red lines) for U10 and V10. However,

this is not the case for T2M and MSLP. For T2M in ED, the RMSEs of YingLong-Pangu and HRRR.F are comparable, while for T2M in WD and MSLP in both domains, the RMSEs of YingLong-Pangu are generally larger than those of HRRR.F.

To test the impact of the LBC provided by the Pangu-weather forecast on the performance of YingLong, the HRRR.F results with 24 km spatial resolution are applied as the LBC for YingLong (denoted as YingLong-HRRR.F24) which provides less information than HRRR.F. Figure 4 shows that the RMSEs of YingLong-HRRR.F24 (blue lines) are generally smaller than those of YingLong-Pangu for all surface variables. Moreover, the RMSEs of YingLong-HRRR.F24 are comparable to those of HRRR.F for T2M and MSLP in WD, and even smaller than those of HRRR.F for T2M in ED. These results indicate that the HRRR.F with 24 km spatial resolution

(a)



(b)

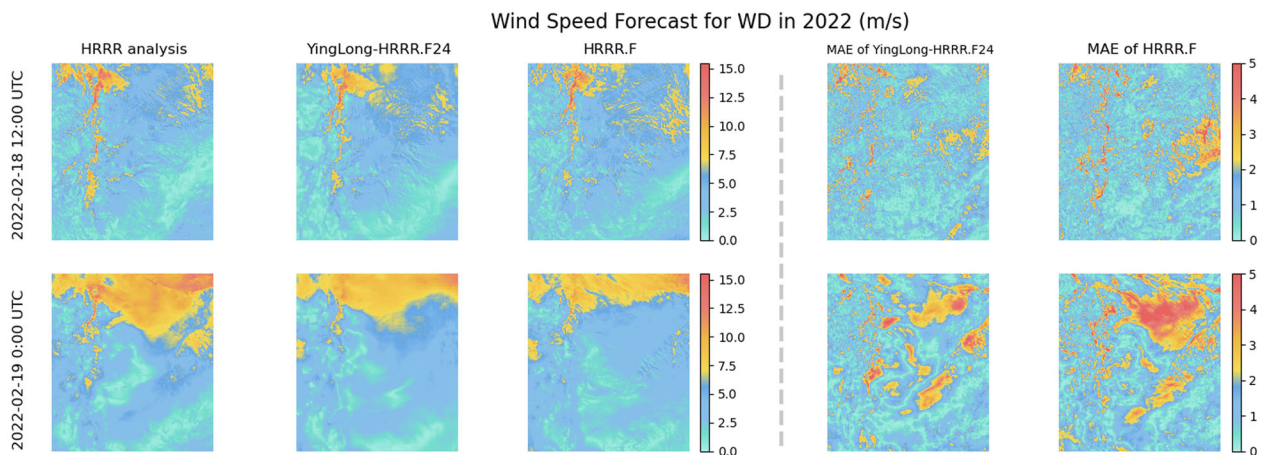


Fig. 5 | Forecast visualization of wind speed forecast results of YingLong model. **a** Forecasts in ED with forecast initialization time at 2022-11-10 00:00 UTC, and lead time is 36 h (first row) and 48 h (second row). The three columns on the left hand side are HRRR.A, YingLong-HRRR.F24, and HRRR.F results at the corresponding

forecast time. The two columns on the right hand side are the spatial distributions of MAE of YingLong-HRRR.F24 and HRRR.F using HRRR.A as benchmark. **b** Forecast visualization of wind speed forecast results in WD are initialized at 2022-02-17 00:00 UTC.

provide a better LBC than the Pangu-weather forecasts. This suggests that the performance of YingLong can be further improved by providing better LBCs. The conclusions drawn from the RMSEs in Fig. 4 are consistent with those from the ACCs (Supplementary Fig. S4), and are further confirmed by the selected visualizations of YingLong-Pangu forecasts in Fig. 5a, b (also see Supplementary Figs. S6–S11 in Supplementary Note 6). Although surface specific humidity observations are available, it is not a variable included in YingLong (see Table 1). Therefore, there are no rolling forecasts for surface specific humidity observations. Finally, the statistical significance test among Yinglong-Pangu, YingLong-HRRR.F24, and HRRR.F (shown in Supplementary Table S1), indicates that the differences between Yinglong models with different LBCs and HRRR.F are statistically significant in most cases for forecasting variables, with p -values less than 0.01 in the negative t -statistic.

To exam the seasonal variability of the predictability, the average values of the cumulative error of the daily 24-h forecast are shown in Supplementary Fig. S12. The figure reveals some interesting features. For example, for all three forecasts of wind speed, the error is general larger in spring (around day 100) and smaller in autumn (around day 200). The T2M (WD) forecast error of Yinglong-Pangu is larger than those of YingLong-HRRR.F24 and HRRR.F in summer (around day 180), indicating that the Pangu forecast is the major error source.

The skills of both YingLong-Pangu and YingLong-HRRR.F24 are higher than those of HRRR.F for predicting U10 and V10, while this is not always the case for predicting T2M and MSLP. The reason could be that U10 and V10 are hydrodynamic variables, while T2M and MSLP are thermodynamic variables that are more sensitive to regional surface conditions, which are not sufficiently described in AI-based models. Specifically, U10 and V10 relate to fluids in motion, which are dominated by pressure gradients. The AI-based YingLong is trained using the analysis states of pressure and wind speed at several geopotential heights (see Table 1). Since the analysis was generated by assimilating observations into HRRR.F, some model errors related to HRRR.F could be corrected by those observations. Therefore, U10 and V10 are likely better predicted by YingLong than by HRRR.F. On the other hand, T2M and MSLP are more closely related to surface radiative forcing. The radiative forcing data is not used for training YingLong but is a forcing variable for HRRR.F. Therefore, T2M and MSLP could be better predicted by HRRR.F than by YingLong. The RMSEs (and ACCs) of all forecasts in ED are smaller (larger) than those in WD. This could be due to the more complex terrain in WD compared to ED.

Accurately predicting wind speed is very important for wind power forecasting, scheduling, operation and maintenance^{41,42}. The wind speed at wind turbine levels (i.e., 85 m) can be estimated from U10 and V10 using a power law, with the power parameter estimated using high-resolution

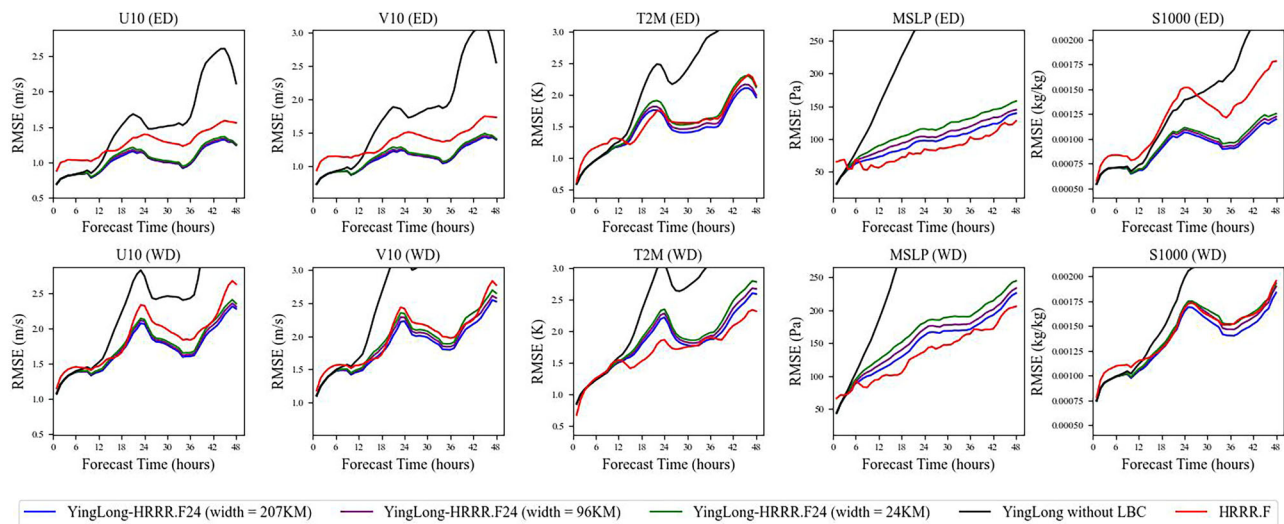


Fig. 6 | The impact of different width of lateral boundary strategies on the forecast effect of inner region. RMSEs (HRRR.A as benchmark) of YingLong-HRRR.F24 forecast in the inner area with no LBC (black lines), lateral boundary with widths 24 km (green lines), 96 km (purple lines), 207 km (blue lines) and HRRR.F (red line).

analysis data or numerical LAM simulations⁴³. Wind data at a 1-hour time scale is a crucial input for modeling the performance and impact of wind farms on many countries' electricity systems⁴³. Since YingLong can forecast hourly high-resolution wind surface speed more promptly and accurately than NWP, it has the potential to be applied to the planning and operation of the wind power generation industry.

Discussions

Results "Forecasting skills of YingLong-Pangu" relates to the forecast skill of YingLong-Pangu, with in-situ observation data as the benchmark. This section addresses modeling issues for improving YingLong, such as the architecture, the LBCs, the homogeneity of training data, and the extremes simulated by YingLong.

The HRRR.A is used as the benchmark for the evaluation and testing of YingLong because it serves as the labels for training YingLong. If other data were used as a benchmark, it would be difficult to determine whether the improvement of a model is due to the model itself or because the model fits the benchmark better. Since the HRRR.A data is a fusion of HRRR.F and observations, HRRR.F with 24 km resolution are chosen as LBC, because they are likely to be more consistent with YingLong forecasts than other LBCs studied in this paper and provide no more information than HRRR.F. Therefore, the modeling issues are less affected by the choice of LBCs. Moreover, the smooth LBC scheme (see Eq. (3)) is imposed, and 207 km is selected as the width of the lateral boundary region. These choices are confirmed as good in the following.

To select a suitable deep neural network architecture, a comprehensive ablation experiment was conducted to compare different model architectures, including the adaptive Fourier neural operator (AFNO⁴⁴, the main module of FourCastNet²¹), the SWIN transformer (the main module of Pangu-weather^{22,45}), and the parallel architecture of AFNO and SWIN (introduced in Methods "Architecture").

Intuitively, the parallel architecture of AFNO and SWIN can effectively extract multi-scale features of meteorological variables. Specifically, AFNO can extract global features of meteorological variables, while SWIN can extract local features. In the parallel architecture, the C channels of the neural network's hidden tensor are divided, with the first $\alpha \cdot C$ channels subjected to the SWIN transformer and the remaining $(1-\alpha) \cdot C$ channels subjected to AFNO. The two calculation results are then concatenated together (see Method "Architecture" for details). This parallel architecture allows for the integration of global and local meteorological variable features, resulting in more accurate predictions. The hyperparameter α is estimated to be 0.25, as determined by the ablation test documented in Method "Training".

YingLong can perform rolling forecasts for the entire area, including the internal area and the lateral boundary region, but only the forecast results in the internal area are the focus of this paper. As the forecast time increases, the analysis variables in the inner area are increasingly affected by the initial values outside the inner area. However, if the information of the variables outside the inner area is not included in the rolling forecast, meaning that YingLong only uses the variables inside the domain as the initial field, then the prediction cannot account for the inner area being affected by the initial conditions outside of the domain in the next step. This could cause the difference between the YingLong forecast and the benchmark to become larger with increasing iterations of the rolling forecasts, as demonstrated in Fig. 6 (black lines). To solve this problem, a lateral boundary region is established outside the inner area. This region contains meteorological information from outside the inner area within a forecast time step Δt of the rolling forecast, so it can affect the next forecast at time $t + \Delta t$ of meteorological variables in the inner area. The LBC can be provided by coarse-resolution global weather forecast models and can be transferred to the inner area through YingLong's forecast on the lateral boundary region (see Methods "Lateral boundary condition").

To obtain high-quality forecast results, it is necessary to ensure that the initial state of the lateral boundary at each step of the rolling forecast closely matches the actual atmospheric state. Directly using the coarse-resolution forecast on the lateral boundaries (coarse LBC, see Methods "Lateral boundary condition") at each step of the rolling forecast may cause discontinuities at the junction between the lateral boundary region and the inner area, potentially affecting the forecast results. However, by employing the boundary smoothing strategy (see Methods "Lateral boundary condition", Eq. (3)), the state of meteorological variables between the lateral boundary region and the inner area can be made more coherent, which could be beneficial for the rolling forecast.

Figure 7 shows the difference in RMSEs (taking the analysis data as the benchmark) when using or not using the boundary smoothing strategy, as well as how this difference changes with the spatial range. It can be seen from Fig. 7 that within the 0–30 km strip from the inner area boundary (the red dashed box in Fig. 1), there is a large difference between the boundary smoothing strategy and the direct replacement strategy. As the area under consideration gets closer to the regional center, such as a strip area 60–90 km from the inner area boundary, the difference between these two boundary strategies is reduced. Within a range greater than 120 km from the inner zone boundary, the difference is negligible. Therefore, different boundary strategies have an impact on AI-LAM forecasts, with a smaller impact closer to the domain center. In addition, the significance tests between YingLong

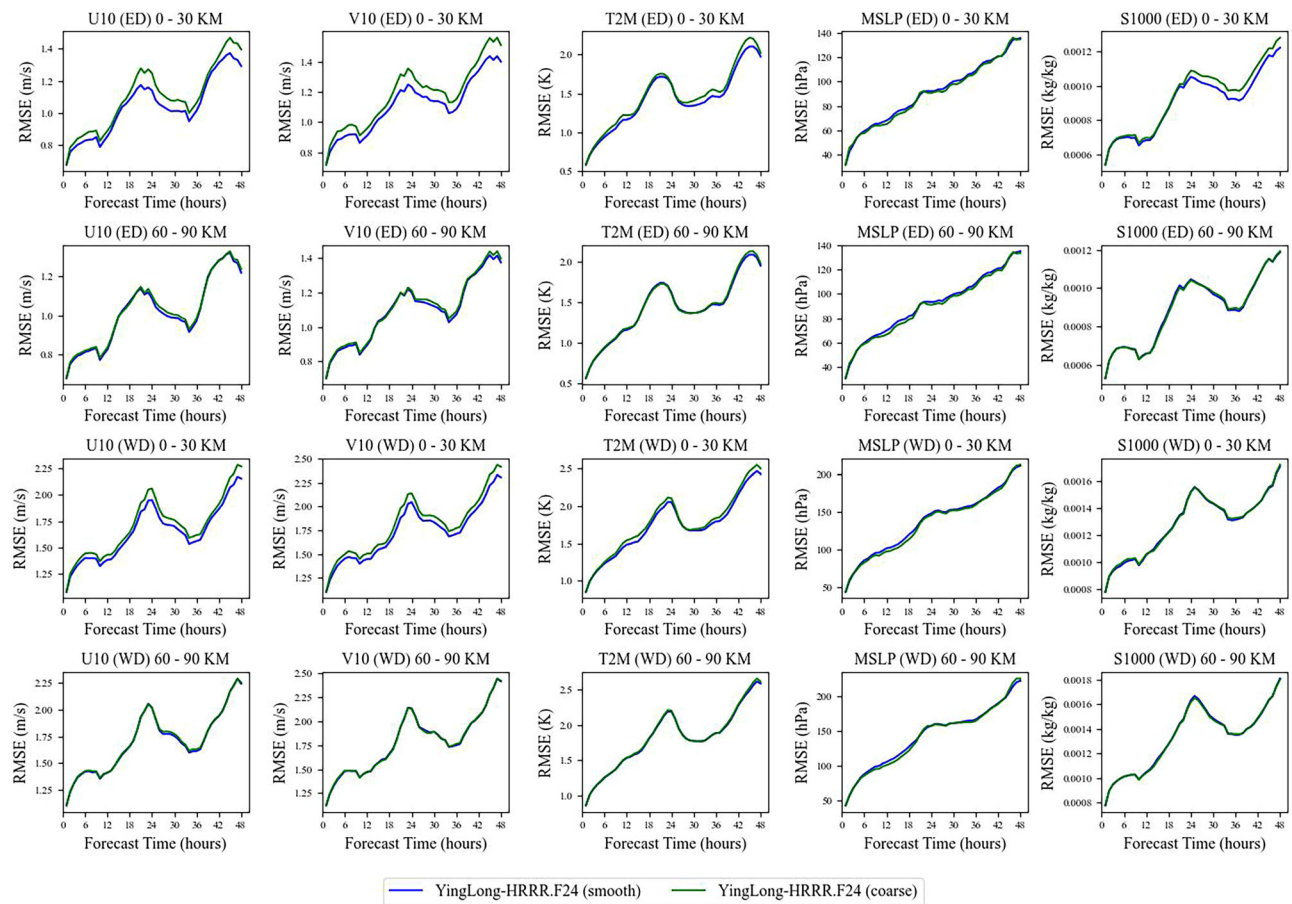


Fig. 7 | The impact of smooth and coarse lateral boundary strategies on the forecast effect at different ranges in the inner region. The RMSEs in this figure are calculated by choosing HRRR.A as benchmark. By selecting strip areas with

distances of 0–30 km and 60–90 km from the inner region boundary, the performance of YingLong-HRRR.F24 with boundary smoothing (blue lines) and coarse (i.e. direct replacement, green lines) are compared.

with coarse and smooth LBC are also made (see Supplementary Section S1.2 of Supplementary Material and Supplementary Table S3 in Supplementary Note 3). The t -statistic is negative in the majority of cases from 12 to 48 h' forecasting, indicating improvement of YingLong with smooth LBC over coarse LBC. The majority of p -values in the negative t -statistic are statistically significant at the 1% level for the strip within the 0–30 km, and at the 5% level for the strip within the 60–90 km.

The width of the lateral boundary region is related to how much information from outside the inner area is introduced into the LAM at each step. A principle for selecting this width is that the analysis data in the inner area at time $t + \Delta t$ should be related to the analysis in the lateral boundary region at time t , but not related to that outside the domain at time t . If the width of the lateral boundary region is too small, the forecast of the inner area at $t + \Delta t$ will lack the necessary information from the lateral boundary region at time t , thus affecting the forecast accuracy for the inner area. Conversely, if the width is too large, some meteorological information in the lateral boundary region that has little to do with the rolling forecast could be introduced, resulting in wasted computational resources. The top speed of level 3 hurricanes in North America is 207 km per hour. Unless there is a level 4 or 5 hurricane, the variability outside of the domain has no influence on the inner area within one hour. To validate the 207 km width against other shorter widths, i.e., 24 km or 96 km, RMSEs of the YingLong-HRRR.F24 forecast in the inner area on smooth LBC with lateral boundary widths of 24 km, 96 km, and 207 km are shown in Fig. 6. The lateral boundary region with a width of 207 km corresponds to the smallest RMSEs. On the other hand, the RMSEs for the lateral boundary region with a width of 96 km are fairly close to those with a width of 207 km, indicating that 207 km is a reasonable choice.

Figure 6 shows that the skills of YingLong-HRRR.F24 are higher than those of HRRR.F for predicting HRRR analysis U10 and V10, while this is not always the case for predicting HRRR analysis T2M and MSLP. This is similar to the skills for predicting station observations (see Fig. 4). Moreover, the skills of YingLong-HRRR.F24 are higher than those of HRRR.F for predicting HRRR analysis S1000, especially in ED, suggesting that YingLong could be more skillful than HRRR.F in forecasting surface specific humidity. Similar conclusion is also supported by ACC results (see Supplementary Fig. S5).

YingLong is constructed on two domains shown in Fig. 1, both with 440×408 grid points in Lambert projection. The training dataset spans the period from 2015 to 2021 and includes HRRR outputs from four different versions of the system (v1–v4). The differences among these versions can be large^{46,47}, including changes in physics schemes, data assimilation methods, and land-surface data. The HRRR datasets for 2022 and 2023, used for testing and validation respectively, are in version v4 only.

To verify the impact of version inhomogeneity, two experiments were performed. The first experiment tested the sensitivity of YingLong's parameters to version inhomogeneity. YingLong was further fine-tuned on HRRR v4 data from 2021 to 2023 and tested on 2024 HRRR data. The RMSEs of the fine-tuned YingLong were very close to those of the original YingLong, indicating that the parameters of YingLong may be robust to version inhomogeneity in this study. The second experiment tested the reliability of only using HRRR v4 data for training. YingLong was trained on HRRR v4 data from 2021 to 2023 and tested on 2024 data. The RMSE of the test result was generally larger than that of YingLong trained using 7-year HRRR v1–v4 data (see Supplementary Fig. S2). This suggests that data length is more important than data homogeneity in this study. Therefore, training

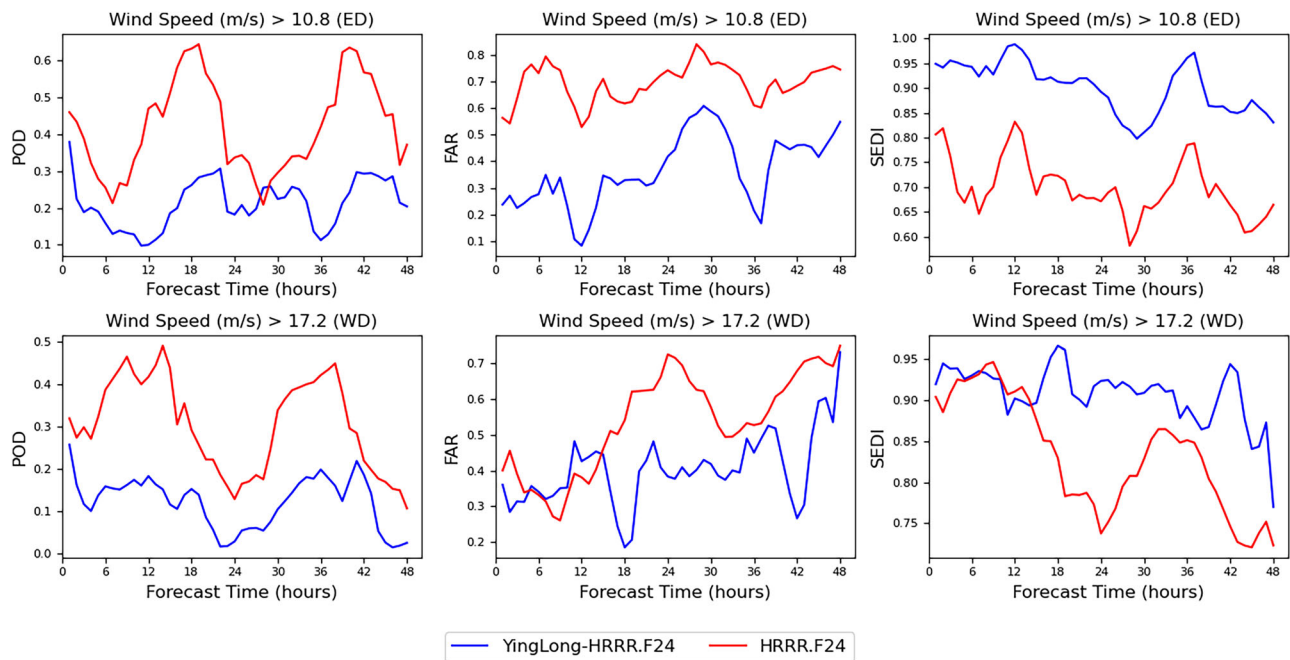


Fig. 8 | Comparisons of POD, FAR and SEDI for wind speed 10 m above surface between YingLong-HRRR.F24 and HRRR.F. From top to bottom, the extreme wind speed thresholds are 10.8 ms^{-1} for ED, and 17.2 ms^{-1} for WD, respectively.

From left to right, the extreme event metrics are POD (higher is better), FAR (lower is better), and SEDI (higher is better). The blue and red lines represent the corresponding results of YingLong and HRRR.F, respectively.

YingLong using 7-year HRRR v1-v4 data is a reasonable choice. The HRRR.A data is hourly, resulting in $7 \times 365 \times 24$ records used for training YingLong, which is roughly comparable to the ERA5 records ($40 \times 365 \times 4$) used for training global AI models.

In this study, wind speed at 10 meters above the surface ($\sqrt{U_{10}^2 + V_{10}^2}$) is selected for evaluating the predictability of extreme wind events for both YingLong and NWP. For WD, the threshold is set at 17.2 ms^{-1} (approximately the Beaufort wind gale scale^{48,49}). Since gale-scale wind is rare for the ED, it is difficult to estimate the three metrics accurately. Therefore, the threshold is reduced to 10.8 ms^{-1} (about a strong breeze). The Probability of Detection (POD, Eq. (S4) in Supplementary Note 4) and False Alarm Ratio (FAR, Eq. (S5) in Supplementary Note 4) of the wind over the two domains are shown in Fig. 8. It shows that both POD and FAR for the YingLong forecast are lower than those of HRRR.F. These facts are consistent with those of global AI-based models⁵⁰. Additionally, Symmetric Extremal Dependence Index (SEDI, which takes into account both the hit rate and the false alarm rate, providing a comprehensive assessment of the prediction accuracy for extreme events, see Eq. (S6) in Supplementary Note 4 for definition) of forecast for wind speed is also shown in Fig. 8. The SEDIs of the YingLong forecast in both domains are larger than those of HRRR.F overall.

Methods

Datasets

The hourly analysis data of the HRRR dataset^{46,47,51} at 3-km resolution, covering the continental US and Alaska, is used for training and evaluating the YingLong model. The HRRR dataset also provides NWP at 3-km resolution, generated through the WRF-ARW (referred to as HRRR.F in this study). The hourly NWP starts from 00UTC and continues for 48 hours.

To test the forecast skill of YingLong-Pangu, the in-situ observation data from HadISD⁴⁰ in 2022 are selected as benchmarks for calculating RMSEs and ACC. The in-situ observation data includes hourly T2M, U10, V10, and MSLP. This study uses a total of 205 observational sites, with 78 sites in WD and 127 sites in ED.

Architecture

YingLong consists of three components shown in Fig. 2: an embedding layer, spatial mixing layers, and a linear decoder. The input data for the

embedding layer are the 24 variables listed in Table 1 and one forcing elevation. They form a tensor with dimensions of $440 \times 408 \times 25$. The embedding layer consists of patch embedding, position embedding and time embedding, to integrate both the spatial and temporal information into the latent tensor. Patch embedding partitions the input tensor into 2805 patches, each patch with a size $8 \times 8 \times 25$. Then, through a convolution layer, each patch is encoded into a C dimensional vector, resulting in the entire input variables being encoded as a tensor of size $2805 \times C$. Position embedding consists of 2805 learnable parameter vectors associated with each of those 2805 relative positions, enabling the model to learn the encoding for each relative position in the region. The dimension of the position embedding vectors is also set to C . Time embedding encodes the specific time information of input data, including year, month, day, and hour, in a vector of dimensions C . Summing up the output vectors from the patch embedding, the position embedding, and the time embedding, yields the output of the embedding layer of size $2805 \times C$. In this study, $C = 768$.

The output of the embedding layer is reshaped into a $55 \times 51 \times C$ tensor and delivered to the spatial mixing layer. The tensor is further split by a ratio α along the channel dimension. The tensor of size $55 \times 51 \times (\alpha \cdot C)$ is delivered to the Local branch, and the remaining tensor of size $55 \times 51 \times ((1-\alpha) \cdot C)$ is sent to the Global branch. The Local and Global branches operate independently and in parallel. The outputs from the two branches are concatenated along the channel dimension, and a new tensor with a dimension of $55 \times 51 \times C$ is obtained. The Local branch contains Window Multi-head Self-Attention (W-MSA) to capture the features within the window, and Shift Window Multi-head Self-Attention (SW-MSA) blocks in the Swin Transformer⁴⁵ to detect possible relationships among different windows by a shifting operation. For W-MSA, padding is performed so that the $55 \times 51 \times (\alpha \cdot C)$ tensor is transferred to $56 \times 56 \times (\alpha \cdot C)$ tensor. Then, it is divided by an 8×8 window to get 7×7 patches. For SW-MSA, the window is shifted by three patches at each time. After a few alternating steps of W-MSA and SW-MSA, the padding data is removed and a tensor of size $55 \times 51 \times (\alpha \cdot C)$ is returned. The Global branch mainly utilizes the Adaptive Fourier Neural Operator (AFNO) block⁴⁴. A 2D fast Fourier transform is applied to the area of 55×51 along each $(1-\alpha) \cdot C$ channel. In the Fourier frequency domain, feature mixing is carried out using a Multilayer Perceptions (MLP) consisting of two fully connected layers. Information from the frequency

Table 2 | Model parameter count for various hyperparameter α values

Model	Number of parameters
YingLong ($\alpha = 0.25$)	60.5 M
SWIN ($\alpha = 1$)	85.0 M
AFNO ($\alpha = 0$)	60.3 M

domain is transferred to the spatial domain by inverse fast Fourier transform. The outputs of the two branches are concatenated along the channel dimension, and a new tensor with a dimension of $55 \times 51 \times C$ is obtained.

An MLP consisting of two fully connected layers is applied to map the channel dimension from C to $(8 \times 8 \times 24)$ for Linear Decoder. The tensor of size $55 \times 51 \times (8 \times 8 \times 24)$ is then reshaped back to $440 \times 408 \times 24$ as the final output of YingLong as the forecast result at $t + \Delta t$ with $\Delta t = 1$ h.

Training

The input for YingLong pretraining, denoted as X_t , is the HRRR analysis in 2015–2021 and the elevation. The output YingLong(X_t) is denoted by $\hat{X}_{t+\Delta t}$, the analysis data $X_{t+\Delta t}$ at the corresponding time is used as labels for supervised learning. The parameters of YingLong are updated by minimizing the loss function

$$\mathcal{L}_1 = \frac{1}{|D_{batch}|} \sum_{t \in D_{batch}} \frac{\|\hat{X}_{t+\Delta t} - X_{t+\Delta t}\|_2}{\|X_{t+\Delta t}\|_2} \quad (1)$$

where the l_2 norm of a vector $X = (x_1, x_2, \dots, x_n)$ is expressed as $\|X\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$, $t \in D_{batch}$ is the initial forecast time in a minibatch⁵² of forecasts in the training dataset. To achieve better predictability at longer forecast time, the pretrained parameters are fine-tuned by minimizing the loss function

$$\mathcal{L}_2 = \frac{1}{|D_{batch}|} \sum_{t \in D_{batch}} \sum_{i=1}^T \frac{\|\hat{X}_{t+i\Delta t} - X_{t+i\Delta t}\|_2}{\|X_{t+i\Delta t}\|_2} \quad (2)$$

where $\hat{X}_{t+(i+1)\Delta t} = \text{YingLong}(\hat{X}_{t+i\Delta t})$ for $i \geq 1$. In this paper, we set $T = 6$. The larger time step i , the more error is induced to $\hat{X}_{t+i\Delta t}$ from outside of the domain. Based on the trial, $T > 6$ not only fails to improve YingLong's long-range forecast skill, but also requires more computational resources.

The Adam optimizer^{52,53} is utilized to minimize the two loss functions. The model employs a cosine learning rate schedule for pre-training, starting at an initial learning rate of 0.005, iterated over 30 epochs. Following pre-training, the model is fine-tuned for 15 epochs using a cosine learning rate schedule with a lower learning rate of 0.0001. YingLong is trained on two A100 Graphical Processing Units, taking approximately 7 days. The trained YingLong takes about 0.5 s to generate a 48-h hourly forecast result on a single A100 Graphical Processing Unit.

In this study, the hyper parameter α is selected from 0 (AFNO), 0.25 (YingLong), 0.5, 0.75, and 1 (SWIN) by conducting ablation tests. For each α , the parameters of YingLong are estimated using HRRR.A in 2015–2021; then the RMSEs of the YingLong-HRRR.F24 forecast with HRRR.A as benchmark in the training period, the validation period 2023 and testing period 2022 are shown in Supplementary Fig. S1 in Supplementary Note 2, respectively. In ED RMSEs for $\alpha = 0.25$ are overall smallest for all the periods. This is also the case in WD, but with less significance to the case of $\alpha = 0$ (AFNO). The number of parameters of the AI-based models are listed in Table 2. It shows that the number of the parameters of YingLong (60.5 M) is less than that of SWIN (85.0 M), so YingLong is trained more economically than SWIN. For all of these reasons, $\alpha = 0.25$ (YingLong) is selected. The statistical significance tests for ablation experiments of Yinglong are shown in Supplementary Table S2. The t -statistic is negative in the majority of cases from different variables from 12 to 48 h' forecasting, indicating

improvement of YingLong over SWIN and AFNO. The majority of p -values in the negative t -statistic are statistically significant at the 5% level for SWIN and the 1% level for AFNO. So YingLong ($\alpha = 0.25$) is validated.

Lateral boundary condition

To run a LAM requires an LBC \hat{Y}_t , which is normally obtained from AI-based global forecast or coarse resolution NWP, given by Pangu-weather or coarsen HRRR to 24 km (i.e. HRRR.F24) depending on the purpose for forecast or research application. Two LBC schemes, including a coarse LBC and a smooth LBC, are tested. For the coarse LBC scheme, a forecast at fine resolution grid n within the lateral boundary region ($\hat{X}_t(n)$) is altered as $\hat{Y}_t(m)$, where m is each coarse resolution grid in the lateral boundary region that is nearest to the fine resolution grid n . For the smooth LBC scheme, it is further smoothed by the YingLong forecast $\hat{X}_t(n)$ based on the inverse distance weighting approach on the lateral boundary. Specifically, for any point n in the lateral boundary, there is a YingLong forecast result $\hat{X}_t(n)$, and the coarse-resolution NWP or AI-based global forecast result $\hat{Y}_t(m)$ at the point m closest to n in the coarse-resolution grid. When the distance from point n to the corresponding inner area, denoted as $d(n)$, is less than the width of the lateral boundary region, which is 207 km in this paper, the result of boundary smoothing $X_t(n)$ (See Fig. 3) can be expressed as

$$X_t(n) = (1 - d(n)/207)\hat{X}_t(n) + d(n)/207\hat{Y}_t(m) \quad (3)$$

When $d(n)$ is greater than the width of the lateral boundary (this is the case in the four corners of lateral boundary area), i.e., 207 km, $X_t(n) = \hat{Y}_t(m)$. The closer the fine resolution grid n to the inner area, the more weight of YingLong forecast at the grid.

Rolling forecast

We take ED as an example to introduce the process of YingLong rolling forecast. Let $X_t(\text{ED})$, $X_t(\text{ED}_{in})$, $X_t(\text{ED}_{out})$ and $X_t(\text{ED}_{LB})$ represent states in entire, inner, outside and lateral boundary regions of ED at time t , respectively. Firstly, YingLong maps an initial state $X_t(\text{ED})$ (the corresponding HRRR.A data) to the forecast state

$$\hat{X}_{t+\Delta t}(\text{ED}) = \text{YingLong}(X_t(\text{ED})) \quad (4)$$

Then $\hat{X}_{t+\Delta t}(\text{ED}_{LB})$ and a coarse resolution LBC $\hat{Y}_{t+\Delta t}(\text{ED}_{LB})$ are smoothed by Eq. (3) to form the initial state $X_{t+\Delta t}(\text{ED})$ for the next time step (see Fig. 3). It should be noted that during the validation and testing phases, the analysis data of the 24 meteorological variables listed in Table 1 at time t and the topography of the corresponding area are first concatenated to form a $440 \times 408 \times 25$ tensor. This tensor is then used as the input for YingLong to obtain the forecast results of these 24 meteorological variables at time $t + \Delta t$, resulting in a $440 \times 408 \times 24$ tensor. YingLong's forecast results cover both the inner domain and the lateral boundary area. In the subsequent steps of the rolling forecast, YingLong's forecast results in the lateral boundary area are smoothed with the corresponding coarse-resolution forecast results. The smoothed boundary results are then spliced with YingLong's forecast results in the inner domain to form the initial field for YingLong's next forecast. Therefore, it can be seen that these 24 meteorological variables are predicted simultaneously.

Data availability

For the training and testing data discussed in this paper, the HRRR dataset is downloaded from <https://rapidrefresh.noaa.gov/hrrr/>, and this website also provides forecasting results of various variables with WRF-ARW. Pangu-weather forecast data is downloaded from <https://weatherbench2.readthedocs.io/en/latest/data-guide.html#forecast-datasets>. The weather station dataset HadISD can be found at <https://www.metoffice.gov.uk/hadobs/hadisd/>. The source data of each figure are available at <https://doi.org/10.5281/zenodo.15233798>.

Code availability

The code architecture of YingLong is developed on PaddlePaddle, a Python-based framework for deep learning, available at <https://github.com/PaddlePaddle/Paddle>. When building the architecture, we utilize part of the Swin transformer (<https://github.com/microsoft/Swin-Transformer>), an AFNO block (<https://github.com/NVlabs/AFNO-transformer>) is also involved. The code for training YingLong is available at <https://github.com/PaddlePaddle/PaddleScience/tree/develop/examples/yinglong>.

Received: 1 February 2024; Accepted: 1 May 2025;

Published online: 15 May 2025

References

- Rothfusz, L. P. et al. Facets: A proposed next-generation paradigm for high-impact weather forecasting. *B. Am. Meteorol. Soc.* **99**, 2025–2043 (2018).
- Alley, R. B., Emanuel, K. A. & Zhang, F. Advances in weather prediction. *Science* **363**, 342–344 (2019).
- Bauer, P., Thorpe, A. & Brunet, G. The quiet revolution of numerical weather prediction. *Nature* **525**, 47–55 (2015).
- Molteni, F., Buizza, R., Palmer, T. N. & Petroliajgis, T. The ECMWF ensemble prediction system: methodology and validation. *Q. J. Roy. Meteor. Soc.* **122**, 73–119 (1996).
- Ritchie, H. et al. Implementation of the semi-Lagrangian method in a high-resolution version of the ECMWF forecast model. *Mon. Weather Rev.* **123**, 489–514 (1995).
- Bauer, P. et al. *The ECMWF scalability programme: progress and plans* (European Centre for Medium Range Weather Forecasts, 2020).
- Schultz, M. G. et al. Can deep learning beat numerical weather prediction?. *Philos. T. R. Soc. A* **379**, 20200097 (2021).
- Balaji, V. Climbing down charney's ladder: machine learning and the post-dennard era of computational climate science. *Philos. T. R. Soc. A* **379**, 20200085 (2021).
- Allen, M. R., Kettleborough, J. & Stainforth, D. Model error in weather and climate forecasting. In: *Seminar on predictability of weather and climate* 279–304 (ECMWF, 2003) <https://www.ecmwf.int/en/elibrary/73407-model-error-weather-and-climate-forecasting>.
- Palmer, T. et al. Representing model uncertainty in weather and climate prediction. *Annu. Rev. Earth Pl. Sc.* **33**, 163–193 (2005).
- Irrgang, C. et al. Towards neural earth system modelling by integrating artificial intelligence in earth system science. *Nat. Mach. Intell.* **3**, 667–674 (2021).
- Reichstein, M. et al. Deep learning and process understanding for data-earth system science. *Nature* **566**, 195–204 (2019).
- Rasp, S. & Thuerey, N. Data-driven medium-range weather prediction with a Resnet pretrained on climate simulations: a new model for weatherbench. *J. Adv. Model. Earth Sy.* **13**, e2020MS002405 (2021).
- Weyn, J. A., Durran, D. R. & Caruana, R. Improving data-driven global weather prediction using deep convolutional neural networks on a cubed sphere. *J. Adv. Model. Earth Sy.* **12**, e2020MS002109 (2020).
- Chen, G. & Wang, W. Short-term precipitation prediction for contiguous United States using deep learning. *Geophys. Res. Lett.* **49**, e2022GL097904 (2022).
- Gao, Z. et al. Earthformer: exploring space-time transformers for earth system forecasting. *Adv. Neural Inf. Process. Syst.* **35**, 25390–25403 (2022).
- Wang, Y., Long, M., Wang, J., Gao, Z. & Yu, P. S. PredRNN: recurrent neural networks for predictive learning using spatiotemporal LSTMs. In *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems* 879–888 (NeurIPS, 2017).
- Rasp, S. et al. WeatherBench: a benchmark data set for data-driven weather forecasting. *J. Adv. Model. Earth Sy.* **12**, e2020MS002203 (2020).
- Rasp, S. et al. WeatherBench2: a benchmark for the next generation of data-driven global weather models. *J. Adv. Model. Earth Sy.* **16**, e2023MS004019 (2024).
- Keisler, R. Forecasting global weather with graph neural networks. Preprint at <https://arxiv.org/pdf/2202.07575> (2022).
- Pathak, J. et al. FourCastNet: a global data-driven high-resolution weather model using adaptive Fourier neural operators. Preprint at <https://arxiv.org/abs/2202.11214> (2022).
- Bi, K. et al. Accurate medium-range global weather forecasting with 3D neural networks. *Nature* **619**, 533–538 (2023).
- Lam, R. et al. Learning skillful medium-range global weather forecasting. *Science* **382**, 1416–1421 (2023).
- Chen, K. et al. FengWu: pushing the skillful global medium-range weather forecast beyond 10 days lead. Preprint at <https://doi.org/10.48550/arXiv.2304.02948> (2023).
- Chen, L. et al. FuXi: a cascade machine learning forecasting system for 15-day global weather forecast. *npj Clim. Atmos. Sci.* **6**, 190 (2023).
- Nguyen, T., Brandstetter, J., Kapoor, A., Gupta, J. K. & Grover, A. Climax: a foundation model for weather and climate. In: *Proceedings of the 40th International Conference on Machine Learning* (PMLR, 2023).
- Bodnar, C. et al. A foundation model for the earth system. Preprint at <https://doi.org/10.48550/arXiv.2405.13063> (2024).
- Lang, S. et al. AIFS-ECMWF's data-driven forecasting system. Preprint at <https://doi.org/10.48550/arXiv.2406.01465> (2024).
- Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. Roy. Meteor. Soc.* **146**, 1999–2049 (2020).
- Haiden, T., Janousek, M., Vitart, F., Ben-Bouallegue, Z. & Prates, F. Evaluation of ECMWF forecasts, including the 2023 upgrade. *ECMWF Technical Memoranda* <https://www.ecmwf.int/en/elibrary/81389-evaluation-ecmwf-forecasts-including-2023-upgrade> (2023).
- Benjamin, S. G. et al. 100 years of progress in forecasting and NWP applications. *Meteorol. Monogr.* **59**, 13–1 (2019).
- Espeholt, L. et al. Deep learning for twelve hour precipitation forecasts. *Nat. Commun.* **13**, 1–10 (2022).
- Andrychowicz, M. et al. Deep learning for day forecasts from sparse observations. Preprint at <https://doi.org/10.48550/arXiv.2306.06079> (2023).
- Oskarsson, J., Landelius, T. & Lindsten, F. Graph-based neural weather prediction for limited area modeling. In: *NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning* (NeurIPS, 2023).
- Pathak, J. et al. Kilometer-scale convection allowing model emulation using generative diffusion modeling. Preprint at <https://doi.org/10.48550/arXiv.2408.10958> (2024).
- Nipen, T. N. et al. Regional data-driven weather modeling with a global stretched-grid. Preprint at <https://doi.org/10.48550/arXiv.2409.02891> (2024).
- Frogner, I.-L., Singleton, A. T., Koltzow, M. Ø & Andrae, U. Convection-permitting ensembles: Challenges related to their design and use. *Q. J. Roy. Meteor. Soc.* **145**, 90–106 (2019).
- Zhao, P. et al. OMG-HD: A high-resolution AI weather model for end-to-end forecasts from observations. Preprint at <https://arxiv.org/pdf/2412.18239> (2024).
- Skamarock, W. et al. A description of the advanced research WRF model version 4. *NCAR Tech. Note NCAR* <https://opensky.ucar.edu/islandora/object/technotes%3A576> (2019).
- Dunn, R. J. H., Willett, K. M., Parker, D. E. & Mitchell, L. Expanding HadISD: quality-controlled, sub-daily station data from 1931. *Geosci. Instrum. Method. Data Syst.* **5**, 473–491 (2016).
- Chen, Y., Wang, Y., Kirschen, D. & Zhang, B. Model-free renewable scenario generation using generative adversarial networks. *IEEE Trans. Power Syst.* **33**, 3265–3275 (2018).
- Zhang, Z. et al. A machine learning model for hub-height short-term wind speed prediction. *Nat. Commun.* **16**, 3195 (2025).
- Turner, R. et al. Creating synthetic wind speed time series for 15 New Zealand wind farms. *J. Appl. Meteorol. Clim.* **50**, 2394–2409 (2011).

44. Guibas, J. et al. Efficient token mixing for transformers via adaptive Fourier neural operators. In: *International Conference on Learning Representations* (ICLR, 2021).
45. Liu, Z. et al. Swin transformer: hierarchical vision transformer using shifted windows. In: *International Conference on Computer Vision* 10012–10022 (IEEE, 2021).
46. Dowell, D. C. et al. The high-resolution rapid refresh (HRRR): an hourly updating convection-allowing forecast model. part I: motivation and system description. *Weather Forecast* **37**, 1371–1395 (2022).
47. James, E. P. et al. The high-resolution rapid refresh (HRRR): an hourly updating convection-allowing forecast model. part II: forecast performance. *Weather Forecast.* **37**, 1397–1417 (2022).
48. Zhong, X. et al. FuXi-Extreme: improving extreme rainfall and wind forecasts with diffusion model. *Sci. China Earth Sci.* **67**, 3696–3708 (2024).
49. Oliver, J. E. *Encyclopedia of word climatology* (Springer, 2005).
50. Bonavita, M. On some limitations of current machine learning weather prediction models. *Geophys. Res. Lett.* **51**, e2023GL107377 (2024).
51. Benjamin, S. G. et al. A north American hourly assimilation and model forecast cycle: The rapid refresh. *Mon. Weather. Rev.* **144**, 1669–1694 (2016).
52. Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. *Dive into deep learning* (Cambridge University Press, 2023).
53. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. In *3rd International Conference on Learning Representations* (ICLR, 2015).

Acknowledgements

This work was supported by Major Program of National Natural Science Foundation of China (NSFC) Nos.12292980, 12292984, NSFC 12405038, NSFC12031016, NSFC12201024, NSFC12426529; National Key R&D Program of China Nos.2023YFA1009000, 2023YFA1009004, 2020YFA0712203, 2020YFA0712201; Beijing Natural Science Foundation BNSF-Z210003; was funded by the Department of Science, Technology and Information of the Ministry of Education, No:8091B042240, and was also Sponsored by CCF-Baidu Open Fund CCF-BAIDU 202317.

Author contributions

P.X., J.Y., and J.Z. designed the project. P.X. and J.Z. managed and supervised the project. P.X., T.G., and Y.W. carried out model design, development, training, and testing. X.G.Z. guided the model's meteorological experiments. P.X., and X.G.Z. wrote the manuscript. X.Z.Z., S.L., and Z.W., helped revise the manuscript and polish the language. Z.Z., X.H., and X.C. deployed the training environment required by the model. All co-authors participated in the discussions and revised the manuscript.

Competing interests

The authors declare no competing interests.

Ethics

We affirm that this research was conducted in accordance with all relevant ethical guidelines and standards.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43247-025-02347-5>.

Correspondence and requests for materials should be addressed to Junping Yin.

Peer review information *Communications Earth & Environment* thanks Mariana Clare and Marina Astitha, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Rahim Barzegar, Joe Aslin and Aliénor Lavergne. [A peer review file is available.]

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025