NUIST-WE: A foundation architecture for weather and energy forecasting integrating meteorological physics and generative pre-training

Haishan Chen,[1,2,*] Ying Liu,[3,4] Haiqi Zhu,[5] Haoxuan Xu,[6] Chenyue Song,[6] Chen Hui,[7] Xianzhu Liu,[7] Siguang Zhu,[1,2] Yi Zhang,[1,2] Xingxing Ma,[1,2] Wei Hu,[7] Bingcheng Wan,[7] Kai Geng,[7] Wenjie Zhang,[8] Lina Xu,[9,10] Qiuji Ding,[1,10] Junwei Wu,[7] Yubao Liu,[1,10] Huan Wu,[1,11] Changming Dong,[1,12] Nengfei Zhu,[3,4] Lei Huang,[3,4] Bitao Xiao,[3,4] Feng Jiang,[1,7,*] and Yuanjian Yang[1,10,*]

[1]Institute of Energy Meteorology, State Key Laboratory of Climate System Prediction and Risk Management, Nanjing University of Information Science and Technology, Nanjing 210044, China
[2]School of Atmospheric Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China
[3]Institute of Energy Meteorology, Guodian Nanjing Automation CO., LTD., Nanjing 210032, China
[4]Weimeide company, Guodian Nanjing Automation CO., LTD., Nanjing 210032, China
[5]School of Medicine and Health, Harbin Institute of Technology, Harbin 150000, China
[6]Faculty of Computing, Harbin Institute of Technology, Harbin 150000, China
[7]School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing 210044, China
[8]School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China
[9]Inner Mongolia Service Center of Meteorology, Hohhot 010051, China
[10]School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing 210044, China
[11]School of Hydrology and Water Resources, Nanjing University of Information Science and Technology, Nanjing 210044, China
[12]School of Marine Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China
*Correspondence: haishan@nuist.edu.cn (H. C.); fjiang@nuist.edu.cn (F. J.); yyj1985@nuist.edu.cn (Y. Y.)

Received: January 16, 2026; Accepted: May 17, 2026; Published Online: May 25, 2026; https://doi.org/10.59717/j.xinn-geo.2026.100232

© 2026 The Author(s). This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

Citation: Chen H., Liu Y., Zhu H., et al. (2026). NUIST-WE: A foundation architecture for weather and energy forecasting integrating meteorological physics and generative pre-training. *The Innovation Geoscience* **4**:100232.

Dear Editor,

The global energy transition is accelerating in response to international climate change mitigation targets.[1] As an inexhaustible clean energy source, wind energy is crucial to future energy mixes,[2,3] accounting for 56% of new renewable energy capacity installation in 2024. By the end of 2024, the global wind power installation capacity reached 1,136 GW, with China accounting for 36.3% of the world's total installed capacity. However, wind power exhibits a stochastic nature and poor storability, which introduces uncertainties in wind power supply. In addition, the inherent dynamic variabilities such as high volatility and spatial heterogeneity of wind energy impose significant challenges to the power system operation. All these issues directly reduce economic benefits and impose operational risks under the scenario of large-scale integration of wind power into the grid system. Therefore, high-precision energy meteorological forecasting plays a critical role in managing modern power systems.

Energy meteorological forecasting fundamentally differs from traditional meteorological forecasting. It requires tight integration of atmospheric dynamic processes such as boundary layer turbulence evolution, synoptic-scale circulation patterns, and terrain-induced flows with the operational characteristics of energy systems, including turbine aerodynamics and wind farm layout effects. This cross-disciplinary nature creates a unique set of requirements: (1) higher-resolution forecasting (with a spatial scale of 1–3 km grid intervals and a time scale of 15–60 minutes) to accurately capture microscale meteorological fluctuations that directly impact turbine performance, (2) physically consistent coupling between meteorological variables and power output to support operational interpretability, (3) scalability across diverse climate zones (temperate, arid, coastal) and site conditions (flatland, mountainous, offshore), and (4) robustness to data scarcity for new wind farms ("cold-start" scenarios).[4]

Current energy meteorological forecasting methods can be roughly categorized into three evolutionary generations, each with inherent limitations in addressing interdisciplinary challenges:

Physics-based models. Relying on numerical weather prediction (NWP) outputs and turbine power curves, these models (e.g., WRF-Solar,[5] mesoscale NWP-engineering hybrid models) prioritize physical interpretability but suffer from two critical limitations: (1) their spatiotemporal resolution fails to capture microscale processes, including wake effects and terrain-induced flow separation,[6] and (2) the decoupling of the meteorological and energy systems modules overlook bidirectional feedback.[7]

Data-driven models. Enabled by big data and deep learning, AI models (e.g., LSTMs, Transformers, graph neural networks) learn statistical patterns from historical meteorological and power data, achieving higher accuracy than physics-based models. However, they face three scalability barriers: (1) site-specific overfitting—models trained on one wind farm perform poorly on others due to neglect of regional meteorological homogeneity, (2) data hunger—requiring 6–12 months of high-quality historical data for reliable training, which is unavailable for new sites, and (3) physical inconsistency—forecasts may violate basic physical and operational constraints, such as unrealistic wind-speed variations or wind speeds exceeding turbine cut-out limits.

General foundation models. Emerging LLM-inspired time-series models (e.g., Time-LLM, Chronos, Timer-XL),[8] these models aim to achieve cross-site generalization through large-scale pre-training. However, they are ill-suited for energy meteorology due to two critical mismatches: (1) domain misalignment—pre-trained on generic time-series without incorporating meteorological physical constraints, leading to 9%–15% higher error than specialized energy forecasting models in energy-specific scenarios,[4] and (2) computational inefficiency: billions of parameters require specialized hardware for inference, which is unavailable in most power grid operation centers. In summary, existing forecasting paradigms fail to adequately address these requirements, resulting in a persistent gap between predictive accuracy and the practical energy system operations.

To address these cross-disciplinary challenges and overcome the trade-offs among accuracy and scalability, we propose the NUIST-WE Foundation Architecture, a novel foundation architecture for energy meteorological forecasting. The dual module framework of meteorological and wind power forecasting large model integrates meteorological physics prior knowledge with generative pre-training, enabling highly accurate, scalable, and physically consistent wind energy forecasting across diverse sites and climate conditions.

THE NUIST FOUNDATION ARCHITECTURE

GLOBAL multi-source data fusion dataset for cross-disciplinary pre-training

This paper constructs a unified pre-training dataset to learn the coupled response of wind power production to multi-scale meteorological conditions and alleviate data scarcity and domain misalignment issues in existing models. The dataset includes:

Meteorological data. The meteorological data include the ERA5 reanalysis dataset from ECMWF at a spatial resolution of 0.25° × 0.25° for 1979–2024, regional WRF (Weather Research and Forecasting) model outputs at 3–5 km

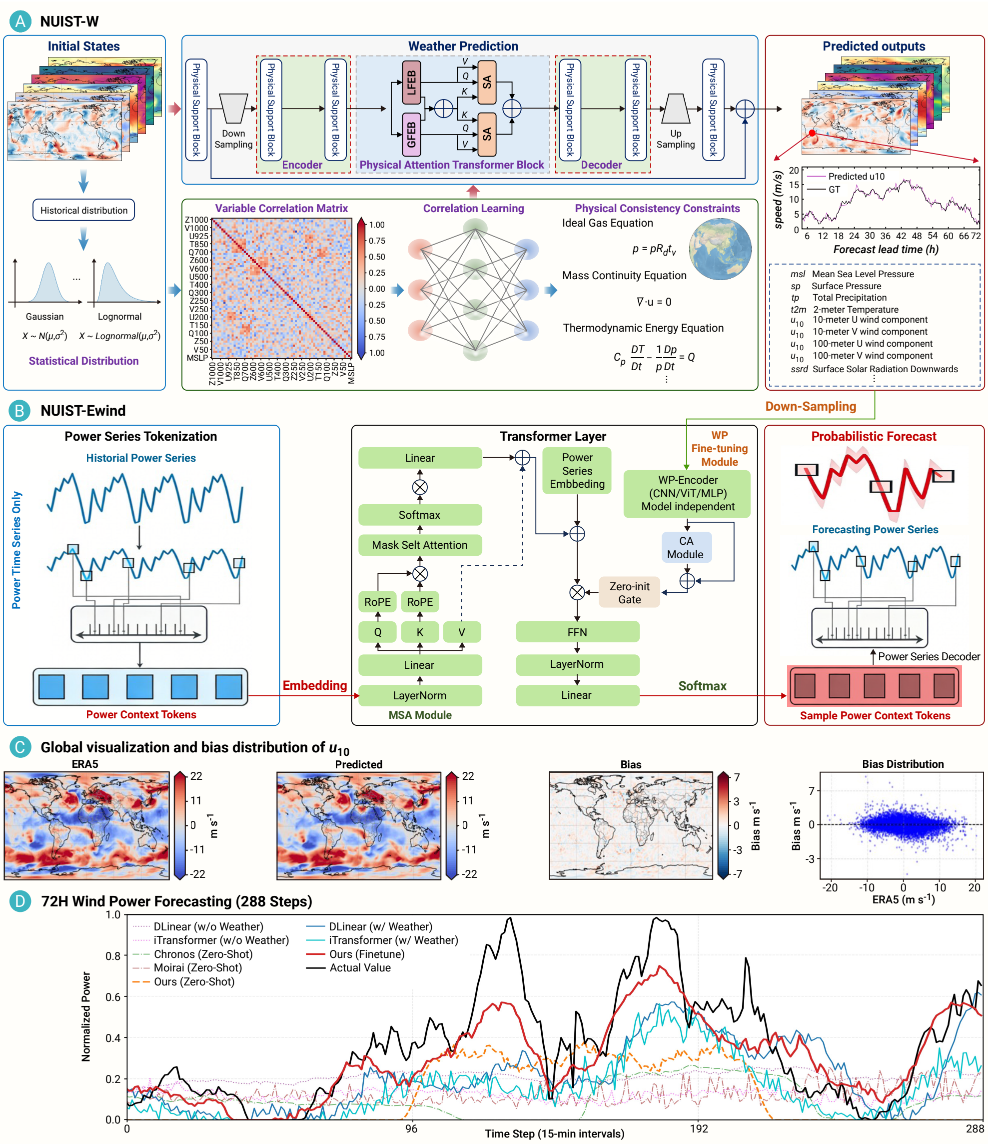

Figure 1. The overall architecture and experiment results of NUIST-WE (A) The architecture of the meteorological base model (NUIST-W) and (B) wind energy forecasting model (NUIST-E_{wind}). (C) and (D) respectively present the comparison between NUIST-W's global visualization and bias distribution of (U10 as an example) and NUIST-E_{wind}'s 72-hour wind power forecasting with the existing mainstream methods for a wind farm in a temperate-flatland area.

resolution, and in situ weather observations from more than 100,000 stations, including wind speed, wind direction, temperature, air pressure, and humidity.

Energy data. Historical power output (at 15-min intervals) from more than 5,000 wind farms globally (2010–2024), covering diverse climate zones (temperate, tropical, arid, polar) and site types (onshore flatland, mountainous, offshore, rooftop solar).

Auxiliary data. Turbine technical parameters (rated power, cut-in/cut-out wind speed), topographic data (DEM, slope, aspect), land cover data (MODIS), and grid operation constraints.

Total dataset volume exceeds 230 billion multivariate time steps, with strict quality control with outlier detection via an IQR method and missing value imputation using physics-informed generative adversarial networks). All meteorological variables and energy data are input in the form of gridded field statistics and represented as 3-D tensors. And the auxiliary data are constraints for meteorological and energy forecasting results.

NUIST-W, a meteorological base model: Climate prior-driven forecasting

NUIST-W (Figure 1A) breaks through the bottlenecks of low resolution and prominent systematic biases, often occurred in conventional NWP models, by fusing meteorological physics prior knowledge with deep learning. It produces refined forecasts of energy-relevant meteorological variables, including wind speed, wind direction, temperature, air pressure, and humidity, at up to 1 km spatial resolution and 15-minute temporal resolution using interpolation-based downscaling. The model enables energy-tailored high-resolution meteorological forecasts through the following mechanisms:

A physical constraint based on climate prior-driven multivariate distribution. To ensure the physical rationality and consistency of long-term forecasting with climate statistics, we compute multi-spatiotemporal statistical features (mean, variance, and quantiles) of key meteorological variables (wind speed, turbulence intensity, and boundary layer height) using historical climate data of 40 years. These features are embedded as physical distribution constraints in the model's latent space via a climate-aware attention mechanism. This mechanism not only modulates forecasting outputs to align with long-term climate patterns but also significantly enhances the model's adaptability across diverse climate zones (e.g., temperate, arid, and coastal regions). For instance, wind speed forecasting for a temperate region is constrained to follow the Weibull distribution parameters (with shape factor 2.0–2.5 and scale factor 6–8 m/s), which is typical for these areas, thereby suppressing physically implausible outliers and improving the physical consistency of the forecasts.

Dual-matrix physical interpretable error correction for the ERA5 data. Targeting systematic biases in ERA5 reanalysis (e.g., wind speed underestimation over mountainous terrain), we design a physically interpretable correction module with (1) a fixed correction matrix derived from historical observations, capturing known biases (e.g., +1.2 m/s wind speed correction for 1000–1500 m elevation), and (2) a low-rank learnable matrix (rank=50) that adapts to unmodeled biases (e.g., seasonal variations in boundary layer stability). This dual-matrix design achieves two-dimensional correction that combines physical consistency and data adaptability, ensuring both the physical interpretability of the correction logic and robust adaptability to complex biases.

NUIST-E_{wind}, a wind energy forecasting model: Generative pre-training for cross-site generalization

NUIST-E_{wind} (Figure 1B) leverages generative pre-training on multi-site wind farm data to learn site-agnostic universal dynamics of energy power sequences. Unlike general time-series foundation models that are plagued by domain misalignment and high computational costs, NUIST-E_{wind} achieves efficient, domain-specific adaptation, addressing the scalability, cross-regional adaptability and cold-start problems in the existing data-driven models.

NUIST-E_{wind} contains a critical component of Energy sequence-specific temporal tokenizer. To adapt LLM generative paradigms to numerical power forecasting, we propose a dynamic vocabulary construction method. Continuous power sequences (0–rated power) are divided into variable-length intervals based on power change rates (e.g., 0.5% rated power per interval for stable periods, and 2% for ramp events), which converts each time-series sequence into a discrete token sequence. This tokenizer preserves critical features (e.g., ramp events, and stable generation periods) while enabling probabilistic forecasting via token sequence generation and multiple sampling methods. The tests on quantization bias and ramp dynamic distortion indicate that the benefits of the adaptive tokenizer far outweigh the minor harms of quantization bias.

Performances of NUIST-W and NUIST-E_{wind}

Superior forecasting performance of the high-resolution output of NUIST-W. NUIST-W effectively captures complex spatio-temporal dependencies among multiple meteorological variables and models nonlinear interactions that are difficult to represent with existing data-driven methods. Real data experiments show that, in comparison to the baseline methods, NUIST-W reduces U-component of 10 m wind speed (U10) root mean square error (RMSE) by 10%–15% (0.15-0.23m/s) (Figure 1C), and 2m temperature RMSE by 12%–16% (0.14–0.19K) respectively, which provides high-quality, high-resolution inputs for the subsequent power forecasting model. Similarly, V-component of 10m wind speed 850 hPa temperature, 500 hPa geopotential height also exhibit lower RMSE with respect to baseline methods.

NWP timeliness-adaptive fine-tuning framework. To address the problem of decreasing NWP reliability with increasing lead-time (e.g., ECMWF NWP wind speed RMSE increases by 0.3 m/s per day), we introduce a learnable gating unit (LGU) that dynamically balances NWP inputs and pre-trained cross-site knowledge. During the fine-tuning process, the LGU adaptively assigns higher weights to NWP (0.7–0.9) for short lead-times (0–6 h) where NWP is accurate, and lower weights (0.3–0.5) for long lead-times (24–72 h) where pre-trained climate patterns are more reliable. For a wind farm in a temperate-flatland area, for example, this adaptive weighting mechanism yields a 7%–10% improvement in forecasting accuracy for lead times > 24 h compared to fixed weight models (Figure 1D). This mechanism learns physically consistent lead-time logic rather than overfitting to specific training distributions. In general, this adaptive weighting mechanism consistently learns robust lead-time-dependent gating logic rather than merely overfitting to training data across diverse climate zones (temperate, arid, and coastal) and terrain conditions (flatland, mountainous, and offshore).

Cross-model adaptable integration. NUIST-E_{wind} features a standardized input interface enabling seamless integration with diverse traditional meteorological model data (e.g., WRF, and ECMWF IFS) and large weather models (GraphCast, Pangu, FourCastNet, etc.). For new wind farms with limited data (cold-start scenarios), NUIST-E_{wind} uses transferable features from pre-training, including recurrent meteorology–power interaction patterns observed across sites and climate regimes, and only requires 1–2 months of historical power data for lightweight adaptation (10–15 GPU hours), compared to 3–6 months for site-specific models. For a test with three new wind farms from different climates and terrains within China for cold-start test scenarios. NUIST-E_{wind} demonstrated stable performance for all three scenarios, with RMSE and MAE reductions of 3%–5% relative to current state-of-the-art methods.

CONCLUSION

This paper introduced a new wind energy forecasting infrastructure that combines meteorological physics with generative pre-training. By embedding domain-specific meteorological priors and physical constraints, NUIST-W enables more accurate and physically consistent meteorological predictions for energy-related applications. NUIST-E_{wind} leverages a dedicated power time-series representation to establish a novel paradigm for energy forecasting. Building upon NUIST-W, the NUIST-E_{wind} architecture is currently being extended to additional energy domains, including hydropower forecasting (NUIST-E_{hydro}), solar thermal energy forecasting (NUIST-E_{solar}), and wave energy forecasting (NUIST-E_{wave}), forming a unified multi-energy forecasting ecosystem.

REFERENCES

1. Zhang Z., Lin L., Gao S., et al. (2025). A machine learning model for hub-height short-term wind speed prediction. *Nat. Commun.* 16:3195. DOI:10.1038/s41467-025-58456-4
2. Dean N. (2022). Collaborating on clean energy action. *Nat. Energy* 7:785–787. DOI:10.1038/s41560-022-01118-2
3. Supran G., Rahmstorf S. and Oreskes N. (2023). Assessing ExxonMobil's global warming projections. *Science* 379:eabk0063. DOI:10.1126/science.abk0063
4. Bodnar C., Bruinsma W. P., Lucic A., et al. (2025). A foundation model for the Earth system. *Nature* 641:1180–1187. DOI:10.1038/s41586-025-09005-y
5. Jimenez P. A., Hacker J. P., Dudhia J., et al. (2016). WRF-Solar: Description and clear-sky assessment of an augmented NWP model for solar power prediction. *Bull. Amer. Meteor. Soc.* 97:1249–1264. DOI:10.1175/BAMS-D-14-00279.1
6. Hersbach H., Bell B., Berrisford P., et al. (2020). The ERA5 global reanalysis. *Q. J. R. Meteorol Soc.* 146:1999–2049. DOI:10.1002/qj.3803
7. Reichstein M., Camps-Valls G., Stevens B., et al. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature* 566:195–204. DOI:10.1038/s41586-019-0912-1
8. Moon J., Kim J., Kang P., et al. (2020). Solving the cold-start problem in short-term load forecasting using tree-based methods. *Energies* 13:886. DOI:10.3390/en13040886

FUNDING AND ACKNOWLEDGMENTS

We would like to thank China Resources New Energy Investment Co., Ltd. Xuzhou Branch and JointErnet Co., Ltd. providing practical application scenarios for this research. Additionally, we are grateful for the computing power support from METAX Computing Power Co., Ltd.. This study was supported by the National Key Research and Development Program of China (Grant number: 2022YFF0801600). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization: Haishan Chen, Feng Jiang, and Yuanjian Yang. Methodology: Haiqi Zhu, Haoxuan Xu, Chenyue Song, Chen Hui, Xianzhu Liu, Kai Geng, and Junwei Wu. Software and model development: Haiqi Zhu, Haoxuan Xu, Chenyue Song, Chen Hui, Xianzhu Liu, Kai Geng, and Junwei Wu. Validation: Haoxuan Xu, Chen Hui, Xianzhu Liu, Kai Geng, and Junwei Wu. Meteorological mechanism analysis: Siguang Zhu, Yi Zhang, Xingxing Ma, and Wei Hu. Data curation: Bingcheng Wan, Wenjie Zhang, Lina Xu, and Qiuji Ding. Resources: Haishan Chen and Ying Liu. Writing - original draft: Haiqi Zhu, Chenyue Song, Haishan Chen, and Xianzhu Liu. Writing - review & editing: Haishan Chen, Yubao Liu, Huan Wu, Changming Dong, Nengfei Zhu, Lei Huang, Bitao Xiao, Feng Jiang, and Yuanjian Yang. Supervision: Haishan Chen, Feng Jiang, and Yuanjian Yang. Project administration: Feng Jiang and Yuanjian Yang. Funding acquisition: Haishan Chen. All authors participated in the review and revision process, jointly contributing to the improvement of the quality of the paper. All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

The authors declare no competing interests.