

Article

Physics-Guided Deep Learning for Short-Term Offshore Wind Power Forecasting

Boni Wang ^{1,2}, Mingjian Zeng ^{1,*}, Haixia Shan ^{3,4,5,6,7,*} , Ming Zhang ¹ and Qiuyu Chen ¹

¹ Jiangsu Meteorological Service Center, Nanjing 210008, China; bnsmiel@163.com (B.W.); z857188112@163.com (M.Z.); xiaoqiuiuaifadai@163.com (Q.C.)

² Key Laboratory of Energy Meteorology, China Meteorological Administration, Beijing 100081, China

³ State Key Laboratory of Climate System Prediction and Risk Management, Nanjing University of Information Science & Technology, Nanjing 210044, China

⁴ International Geophysical Fluid Research Center, School of Marine Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China

⁵ Oceanic Modeling and Observation Laboratory, School of Marine Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China

⁶ School of Marine Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China

⁷ Southern Ocean Science and Engineering Guangdong Laboratory, Zhuhai 519000, China

* Correspondence: jsqxservic@163.com (M.Z.); karen0009@nuist.edu.cn (H.S.)

Abstract

Given wind energy's growing significance in the world's energy structure, the demand for high-precision forecasting is more urgent than ever. However, wind power's inherent non-stationarity is linked to complex and variable meteorological conditions, which pose significant challenges for accurate forecasting. The accuracy of short-term wind power forecasts hinges on estimated future wind speed. Systematic biases often degrade the accuracy of Weather Research and Forecasting (WRF) forecasts. A hybrid LSTM-LightGBM correction model is proposed to correct the WRF wind speed bias. The wind correction model significantly reduces the systematic wind speed bias in the WRF model and achieves a notable reduction in RMSE across the entire wind speed range, with the highest RMSE decreasing by 7%. A new physics-guided and deep learning-integrated model is designed for 24 h short-term wind power forecasting, which integrates physical laws with data-driven features, effectively alleviating the “black-box” problem of purely data-driven models and making prediction results more consistent with engineering practice. The model's predicted wind power has an MAE of 103.94 kW and an R^2 of 0.9853, demonstrating superior prediction capability. The results provide reliable technical support for the real-time scheduling of wind farms and support the reliable, stable operation of the power system.

Keywords: wind power forecasting; short-term prediction; deep learning; permutation feature importance



Academic Editor: Eugen Rusu

Received: 9 June 2026

Revised: 8 July 2026

Accepted: 9 July 2026

Published: 14 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Wind energy is a highly promising source for large-scale development and commercialization. In the context of accelerating the global energy transformation toward green and low-carbon development, wind energy has gradually become a major pillar of electricity supply in a new-type power system. Conversely, the inherent instability, uncertainty, and volatility of wind energy bring great challenges for the secure and stable operation of power systems [1]. Precise wind power prediction can be applied for careful scheduling

of equipment maintenance, efficient utilization of conventional electricity generation, and improvement of overall reliability of the power grid [2].

The precision of short-term wind power forecasting relies on the accurate estimation of future wind speed [3]. Currently, Numerical Weather Prediction (NWP) models, which numerically resolve the fundamental equations of atmospheric motion to provide forecasts of meteorological element fields from hours to days ahead [4], have become the primary data source for short-term wind power prediction [5–8]. Given its open-source availability, flexible configuration, and strong support for high-resolution nested numerical simulations, the WRF model has been widely adopted in the practical wind speed forecasting operations [9]. A conventional technical strategy integrates the high spatiotemporal resolution wind speed outputs from WRF with turbine power curves or data-driven mapping models to predict future wind power.

However, wind speed forecasts from the WRF model inevitably contain errors. These errors are primarily from the following sources: uncertainties from the initial and boundary conditions derived from the global models [10]; limitations in the approximate representation of sub-grid-scale processes by physical parameterization schemes (e.g., planetary boundary layer, microphysical processes, cumulus convection parameterization) [11]; the limited model resolution, which struggles to fully resolve complex terrain and surface heterogeneity [12]; and interpolation biases introduced when extrapolating model-level winds to the turbine hub height [13]. Since the wind power varies approximately as the cube of wind speed, errors in wind speed are nonlinearly magnified in wind power predictions, making the precision of wind speed predictions a key bottleneck in prediction accuracy.

There has been a large amount of research concerning statistical/machine learning techniques for enhancing the precision of wind speed predictions from numerical weather models. Model Output Statistics and Kalman Filtering is the early statistical methods, which both establish linear statistical relationships to correct for systematic bias; they do, however, have limited capability in capturing nonlinear features of the data [14]. Other machine learning-based techniques such as Support Vector Machines (SVMs) [15], Random Forests (RFs) [16], XGBoost [17] and LightGBM [18] have had great success in improving the accuracy of temperature forecasts made with numeric model simulations. Recently, deep learning technologies have demonstrated their capacity to capture sequential dependencies. Recurrent neural networks (RNNs) [19], Gated Recurrent Unit (GRU) [20], Long Short-Term Memory (LSTM) [21], and Bidirectional Long Short-Term Memory (BiLSTM) [5] serve as the main architectures for their capacity to learn long-term dependencies within time series, and the introduction of attention mechanisms has improved the performance of corrections [22]. Given that wind speed data exhibit non-stationary behavior, signal decomposition techniques are commonly applied at the stage of data preprocessing. These techniques include Empirical Mode Decomposition (EMD) [23], Ensemble Empirical Mode Decomposition (EEMD) [24], Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) [25], and Variational Mode Decomposition (VMD) [26]. Some advanced hybrid frameworks use a “decomposition-clustering-matching” approach [3]. This involves breaking down the error series using CEEMDAN or VMD. After this, deep learning models are used to predict the trend component. At the same time, the non-stationary fluctuation component is analyzed using a combination of feature extraction and cluster analysis [27].

The Wind power prediction models have also transitioned from using basic ML techniques (such as Artificial Neural Networks (ANNs) [28] and Support Vector Regression (SVR) [29] to using more sophisticated algorithms (e.g., Hybrid Convolutional Neural Network—Long-Short Term Memory networks (CNN-LSTM) [30]), which can represent both spatial and temporal dependencies and then to using the latest Graph Neural Net-

works (GNNs) algorithm [31], which can model spatial topological relationships. The Transformer model [32] with strong global dependency modeling capabilities has been employed in wind power forecasting. Currently, hybrid models that incorporate signal decomposition, feature engineering, optimization techniques and deep learning have become mainstream methods.

Although substantial progress has been made in research on the correction of wind speeds and the forecasting of wind power, there are still certain limitations in the current studies. Most data-driven models rely heavily on manual feature engineering [33], and the models themselves are often “black box” structures, failing to fully integrate the physical mechanisms that cause forecast errors [34]. These issues can result in biases in wind power forecasting, and the generalizability and robustness of the model still require further enhancement. To tackle these challenges, this study first develops a hybrid artificial intelligence model for wind speed correction. The model integrates the advantages of LSTM and LightGBM, while introducing physical constraints and smoothing strategies to effectively mitigate systematic biases in the WRF model. Then, the corrected wind speed series are input into the physics-guided deep learning model for wind power forecasting. The model incorporates physically meaningful variables derived from the physics of wind power generation, rather than solely relying on conventional data-driven inputs, to enhance predictive performance and physical interpretability. In addition, we conduct comprehensive comparative experiments and ablation studies to evaluate the contribution of each component in the model. In the meantime, the permutation feature importance and physical coefficient uncertainty analysis are employed to explore the internal learning behavior of the proposed framework and evaluate its reliability from the perspective of physical consistency.

2. Materials and Methods

2.1. Data and Preprocessing

To validate the performance of the proposed model, a dataset of wind speed observations with 15 min intervals and the corresponding wind power output from a wind farm located in Jiangsu Province, China (121.6033° E, 32.6467° N) was collected in 2024. The wind turbines of this farm have a rated power of 4000 kW, a hub height of 100 m, a cut-in wind speed of 4 m/s, a rated wind speed of 12 m/s, and a cut-out wind speed of 25 m/s. Data loss often happens during the operational phase of wind turbines, primarily due to equipment failures, maintenance activities, and other related reasons. Meanwhile, environmental parameters such as temperature, humidity, and atmospheric pressure, and the status of the equipment can also cause deviation of the wind speed-power conversion characteristic of the turbines from the ideal curve. The accuracy and integrity of data are the basis for improving the reliability of wind power prediction. Therefore, this study conducts preprocessing on wind speed and wind power data to effectively eliminate abnormal data.

The implementation rules (Equation (1)) for data cleaning are as follows: wind power P is set to zero for wind speeds below cut-in or above cut-out. For wind speeds between cut-in and rated, wind power remains unchanged. Between rated and cut-out, wind power is capped at the rated level, representing full load operation. The cleaned wind speed–wind power curve is shown in Figure 1. There are some clearly abnormal data points: for instance, when the wind speed is between 5 and 15 m/s, some power values are lower than expected and clearly deviate from the theoretical curve. The interquartile range (IQR) method is also

adopted in this study to identify abnormal data. The principle of the data cleaning method is graphically illustrated in Figure 2.

$$P = \begin{cases} P & u_{cut-in} < u < u_{rated} \\ P_{rated} & u_{rated} \leq u < u_{cut-out} \\ 0 & u > u_{cut-out} \text{ or } 0 < u < u_{cut-in} \end{cases} \quad (1)$$

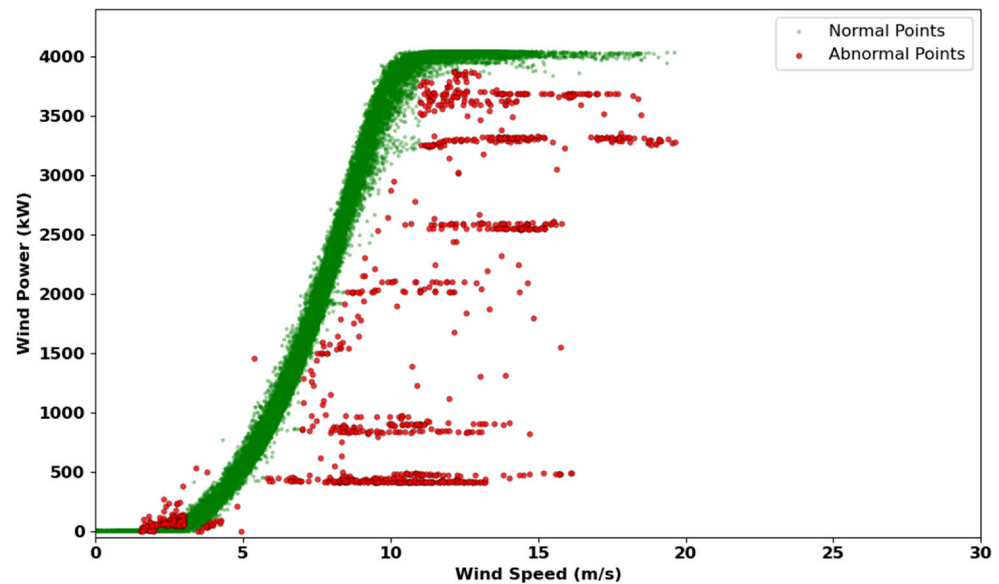


Figure 1. Distribution of normal (green) and abnormal (red) points identified by the Interquartile Range (IQR) Method.

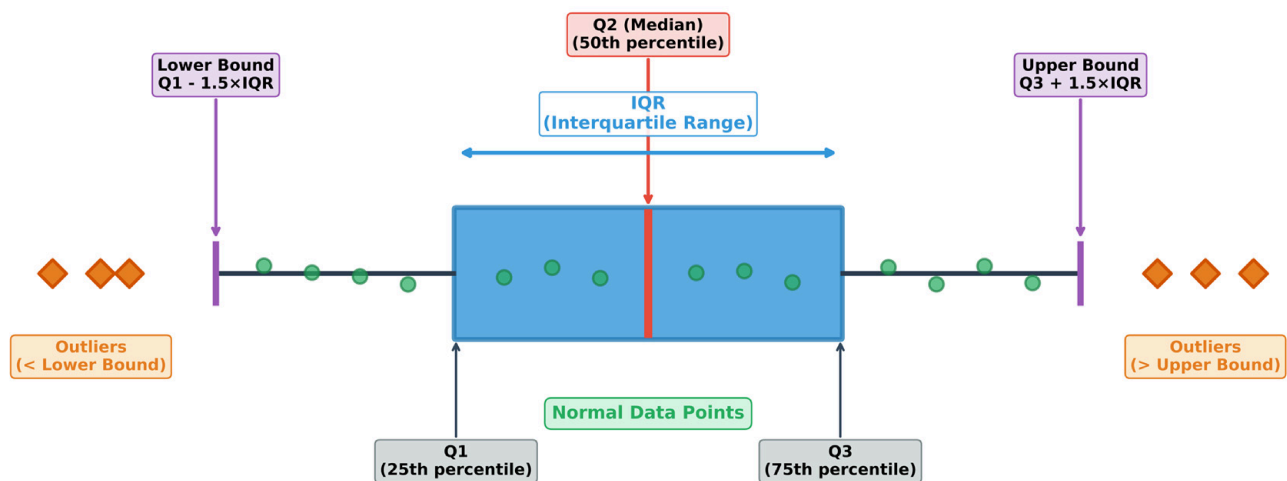


Figure 2. Schematic diagram of the quartile method for data cleaning (not to scale).

All data preprocessing, feature calculation and subsequent deep learning modeling algorithms in this paper are fully implemented based on Python 3.10.19. Core libraries including NumPy 2.2.6, Pandas 2.3.3, Scikit-learn 1.7.1, Matplotlib 3.10.8 and Seaborn 0.13.2 are used for data manipulation and statistical visualization, while TensorFlow 2.10 with Keras API is adopted to build and train the physics-informed prediction model. All numerical experiments are conducted on a workstation equipped with an NVIDIA RTX 4080 GPU, and fixed random seeds are configured in the code to ensure experimental reproducibility.

Unlike the Z-score method [35], which is computed from the mean and standard deviation, the IQR method relies on the quartiles of the data [36]. Because it is not affected

by extreme values, this method is more robust. Specifically, for each wind speed range, the first quartile (Q1, 25th percentile) and third quartile (Q3, 75th percentile) of the wind power values are calculated, and the IQR is then found using the formula $IQR = Q3 - Q1$. Subsequently, the range of valid data is defined as $[Q1 - k \cdot IQR, Q3 + k \cdot IQR]$. To maintain rigorous data preprocessing criteria, a coefficient of $k = 1.5$ is utilized in this research.

This method is particularly suitable for processing datasets that may have a skewed distribution or be disturbed by potential outliers. With an overall outlier rate of only 3.18%, it indicates that the original data is of high quality. Moreover, the cleaning rules not only accurately eliminate outliers but also avoid data loss caused by over-screening, thereby providing high-quality data for subsequent wind speed correction and wind power prediction.

2.2. Wind Speed Correction Model

The WRF model [9] is a numerical weather prediction model for mesoscale weather systems, designed for simulation and real-time forecasting. The WRF model is entirely compressible and non-hydrostatic. The Arakawa C-grid serves as its grid staggering, and the vertical coordinate is a hydrostatic pressure coordinate that follows the topography. The Runge–Kutta 2nd and 3rd order temporal integration schemes, as well as 2nd to 6th order advection schemes, are used in both horizontal and vertical directions by the WRF model. As a popular NWP model, the WRF model is appropriate for weather studies spanning from a few meters to thousands of kilometers. Thus, the current study used the WRF model to estimate wind speed as an input to the wind power model.

This study proposes a hybrid wind speed correction model (shown in Figure 3) based on a combination of LSTM and LightGBM, aiming to correct biases in wind speed outputs from the WRF model. The correction model first takes 100 m WRF forecasts of the wind speed, air temperature, relative humidity, and pressure, together with the wind speed forecast error from the previous time step, as input. The target is the WRF wind speed forecast error at the current time step. After all the features are normalized using Min-MaxScaler, they are input into a double-layer LSTM network with a sliding window of 96-time steps (corresponding to 15 min interval data over 24 h). The two-layer LSTM contains 64 and 32 hidden units, respectively, both employing the tanh activation function, followed by a 20% dropout layer to prevent overfitting. The main advantage of LSTM lies in that it can selectively memorize or discard historical information, thereby effectively capturing long-range temporal dependencies in meteorological variables. This property is particularly suitable for wind speed forecast error correction, as forecast errors typically exhibit persistence associated with the evolution of weather systems. The preliminary corrected wind speed is derived by incorporating the wind speed error predicted by LSTM to the original WRF forecast.

The refined wind speed correction is achieved by feeding a comprehensive set of features into a LightGBM model. The basic features include the LSTM-corrected wind speed, the original WRF wind speed, and their difference, along with the WRF forecasts of air temperature, relative humidity, and pressure. Temporal features such as cosine/sine components of forecast time, day of the week, and month are also included to capture diurnal and seasonal patterns. Lagged features cover residuals and LSTM-corrected wind speeds at lags of 1, 2, 3, 4, 6, 12, and 24-time steps. Rolling statistics include the mean and standard deviation of residuals, as well as the mean of LSTM-corrected values, calculated over window sizes of 4, 12, 24, and 48-time steps. Difference features include first- and fourth-order differences in the LSTM-corrected wind speed, along with the first-order difference in the raw WRF wind speed forecast. The interaction features constructed in this study encompass air temperature, humidity, the wind speed-to-temperature ratio, and

atmospheric pressure anomalies. Additionally, a residual trend feature is introduced to characterize long-term directional variations in the underlying data.

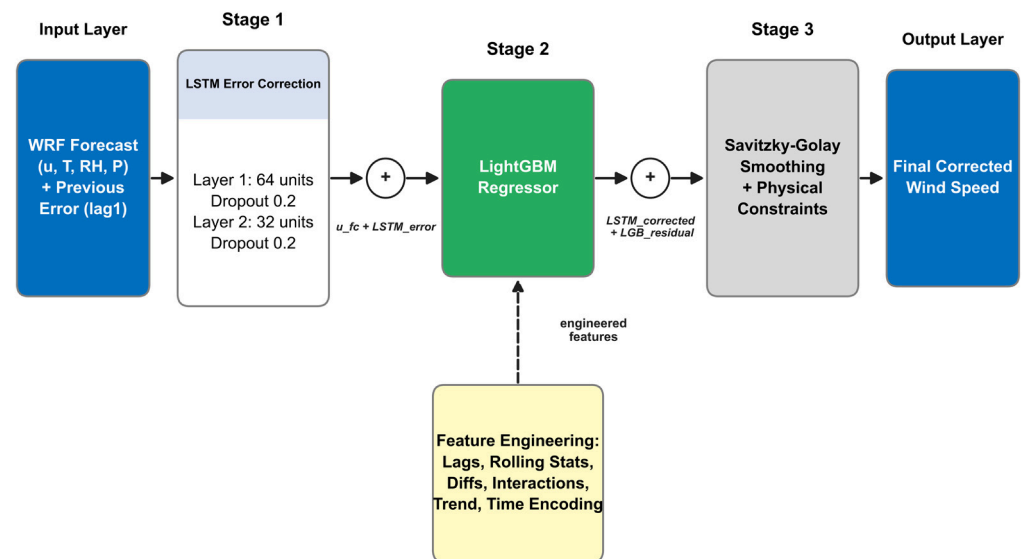


Figure 3. The framework for the hybrid LSTM–LightGBM wind speed correction model.

A Savitzky–Golay (SG) filter [37] with a window size of 11-time steps and a second-order polynomial is employed to smooth the outputs of the LightGBM model. By performing local polynomial least-squares fitting, the SG filter effectively attenuates high-frequency noise while retaining the peak characteristics of wind speed variations. Concurrently, a non-negativity constraint is introduced to guarantee the physical rationality of the corrected wind speed, thereby generating the final refined wind speed series.

2.3. Wind Power Model

Wind power generation essentially constitutes a conversion process from wind energy to mechanical energy and subsequently to electrical energy. The wind power available at the turbine aerodynamics level flowing through the rotor swept area is usually expressed based on the actuator-disk theory and the Betz-Joukowski theory [38,39]. The corresponding expression can be expressed as follows:

$$P_{wind} = \frac{1}{2} \rho u^3 A \quad (2)$$

where ρ is the air density (kg/m^3), u is wind speed (m/s), and A is the rotor swept area (m^2). But only a fraction of the available wind power can be turned into useful output from turbines. The actual aerodynamic power extracted by the turbine is thus usually characterized by introducing the power coefficient C_p [38,39]:

$$P = \frac{1}{2} \rho u^3 A C_p \quad (3)$$

where C_p is the aerodynamic conversion efficiency that depends on turbine operating and design conditions, in particular the tip-speed ratio and blade pitch angle [39].

The wind power forecasting model further relates air density to easily obtainable meteorological parameters using the ideal gas law, forming a computable physical feature term. When wind speed increases sharply (positive ramp), the pitch system of the wind turbine will act to protect the generator, resulting in a delayed response in the growth of actual power. By introducing the power ramp rate $\Delta(\rho u^3)$, the model can learn the

dynamic response of such a physical system. Additionally, the rate of change in wind speed (acceleration) influences wind power response speed due to rotor inertia [40], for which a corresponding correction term $\frac{du}{dt}$ is introduced. Ultimately, a learnable physical correction module is established. The model's final prediction output consists of a deep learning base prediction and a physics-informed correction, represented mathematically as follows:

$$\hat{P} = P_{base} + (w_1 \rho u^3 + w_2 \Delta(\rho u^3) + w_3 \frac{du}{dt}) \quad (4)$$

In Equation (4), P_{base} represents the base power prediction from the deep learning component. The variables w_1 , w_2 , and w_3 are adjustable weights related to physical characteristics, which serve to regulate the impact of physical priors in relation to data-driven features, thereby mitigating potential prediction inaccuracies from the failure of physical assumptions during periods of extreme weather. For the parameter initialization, we set the initial values of w_1 , w_2 , and w_3 to a constant of 0.8. The optimization is initiated from a fixed starting point, ensuring the optimization starts from a physically plausible range consistent with the wind farm aerodynamic characteristics. During model training, all three learnable weights are constrained by non-negative kernel constraints to meet strict physical boundary conditions. Moreover, the optimizer employs gradient clipping with a norm threshold of 0.5 to prevent excessively large parameter updates, which results in more stable convergence of the physical coefficients. We train an ensemble of five models with different random seeds to test the robustness of the learned physical weights. This integrated approach guarantees the model's physical soundness under typical operating conditions while preserving the adaptability of data-driven techniques for atypical scenarios.

A physics-guided and deep learning-integrated wind power forecasting model is proposed in this study, which is based on a dual-branch encoding–decoding correction framework (Figure 4). This model can achieve efficient multi-scale information utilization through the parallel processing and cross-fusion of historical and future sequence branches. The historical sequence branch processes the past 72 h data (observed wind speed, wind power, cubic wind speed, wind speed change, wind power change, and the standard deviation of wind speed over the past four time steps). A masking layer is first used to filter out the interference of missing values. A two-layer bidirectional GRU network is then employed. GRU simplifies the gating structure compared to LSTM, reducing computation while retaining the capability to learn long-term temporal dependencies. The bidirectional design simultaneously captures trend changes from the past to the present and from the present back to the past. A fully connected layer will compress the features to 64 dimensions for subsequent fusion. The future sequence branch processes the future 24 h NWP data (corrected wind speed, forecast temperature, pressure, humidity, forecasting hour encoded using sine/cosine functions, the intersection term between wind speed and the cosine component of forecast time) and can be divided into a CNN sub-branch and a Transformer sub-branch. In the CNN sub-branch, two 1D convolutional layers with kernels of size 5 and 3 are used, together with batch normalization, swish activation [41], and residual connections. A Squeeze-and-Excitation (SE) attention module [42] is incorporated for channel-wise feature recalibration. The Transformer path captures long-range temporal dependencies through 4-head multi-head self-attention and feed-forward modules. In the feature fusion and decoding stage, the features from the historical branch, the CNN sub-branch, and the Transformer sub-branch are concatenated and fed into a bidirectional GRU decoding layer. As an innovation of the correction model, the physical correction module extracts the density power, wind acceleration, and wind power ramp rate from the future sequence branch. Following weight learning via three independent fully connected layers,

the learned features are finally fused with the base prediction using a weight coefficient of 0.1 to yield the final forecasting result.

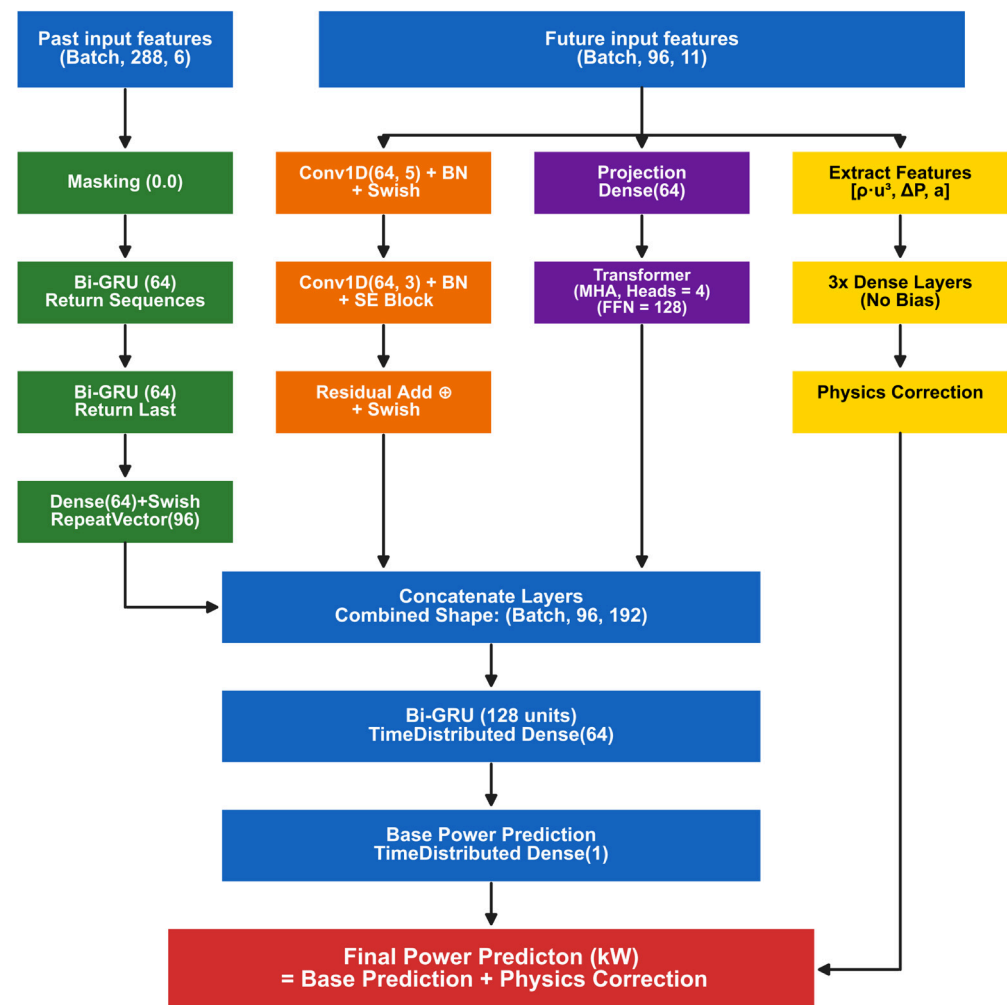


Figure 4. The framework for the physics-guided and deep learning-integrated wind power forecasting.

A weighted composite loss function (Equation (5)) is utilized in this study, which consists of an 80% Mean Squared Error (MSE) term for accurate prediction, a 10% logcosh term for outlier resistance, and a 10% smoothness constraint for preserving the temporal continuity of prediction outputs.

$$\begin{aligned}
 Loss = & 0.8 \times \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 + 0.1 \times \frac{1}{N} \sum_{i=1}^N \log(\cosh(\hat{y}_i - y_i)) \\
 & + 0.1 \times \frac{1}{N-1} \sum_{i=2}^N |(\hat{y}_i - \hat{y}_{i-1}) - (y_i - y_{i-1})|
 \end{aligned} \quad (5)$$

where y_i denotes the actual wind power, $\hat{y}_i$ signifies the predicted wind power, and N represents the sample size.

During the data preprocessing stage, historical observation data is temporally aligned and duplicate removed. All timestamp information is standardized to UTC. Historical sequential data is preprocessed with the RobustScaler normalization [43]. This scaling method uses the median and IQR to transform the data, which reduces the adverse effect of the outliers, reduces the sensitivity of the model to extreme values, and helps the normalized sequences maintain the intrinsic distribution characteristics of the original data. The target wind power data are normalized with StandardScaler according to the data

mean and variance, aiming to facilitate the calculation of the loss function and accelerate the convergence of the model. We split the dataset into training and testing sets with an 8:2 ratio and a batch size of 24. We propose an ensemble learning approach to improve model generalization and reduce the prediction variance. Five models are trained independently by using different random seeds (42, 101, 777, 2024, 999), and their outputs are ensemble to get the final prediction. For model optimization, the Adam optimizer is integrated with a cosine annealing warm restart learning rate strategy, with an initial learning rate of 10^{-3} and the first decay point set to one-third of the total training steps. Additionally, an early stopping mechanism is employed to prevent overfitting, and the training will be halted if the validation loss does not improve for 30 consecutive epochs. To ensure physical consistency, a post-processing mask has been applied after the model output: whenever the input wind speed is below the cut-in speed or above the cut-out speed, the corresponding power prediction is forced to zero.

2.4. Evaluation Metrics

The performance of the proposed wind power forecasting model is quantitatively assessed by three widely used metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2). MAE is the average of absolute error between predicted and observed data. It is easy to interpret as the average error of a model in terms of prediction. It is not sensitive to outliers and thus is appropriate for evaluating the average prediction bias of the model. RMSE tells about the spread of the prediction errors. The power of two magnifies large deviations. This allows us to effectively assess the robustness of a model. As for R^2 , it describes the explanatory power of the fitted model for data variation, and values closer to 1 indicate better forecasting performance. All the above metrics are calculated after applying inverse transforming the normalized prediction results to ensure a clear physical interpretation.

$$MAE = \sqrt{\frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|} \quad (6)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (7)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (8)$$

3. Results

3.1. Study Area and WRF Initialization

Figure 5 shows the location of the selected wind farm and the model domain. Daily WRF prediction is initiated at 00:00 UTC, providing 24 h forecasts at a 15 min temporal resolution. The initialization and boundary conditions of the WRF prediction are 0.25 output datasets of the Global Forecast System Model, and topographic data uses static geographic data of the WRF Preprocessing System. The WRF model is configured with two nested grids at horizontal resolutions of 9 and 3 km. The WRF simulation used the Lambert projection, and the vertical layer is 38.

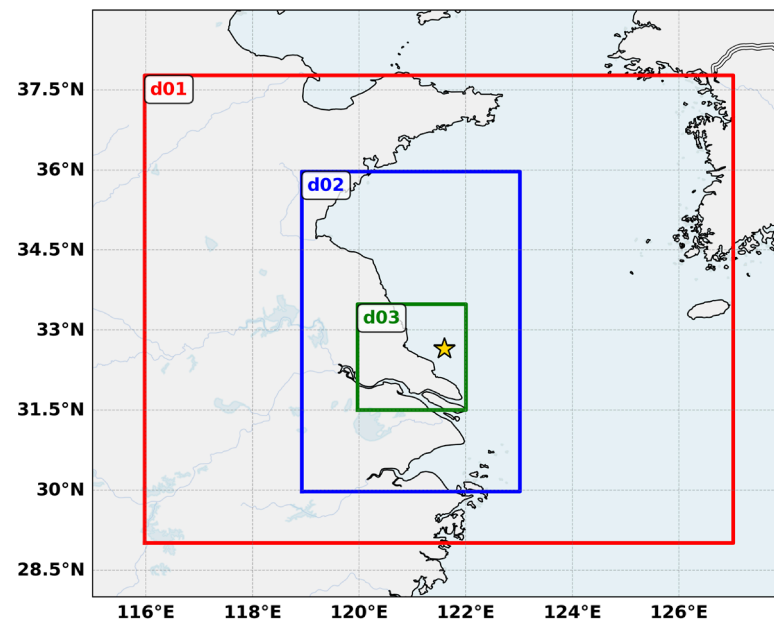


Figure 5. The Schematic diagram of the nesting WRF model. Domain D01 (red box) ranges from 115.98° E to 127.02° E in longitude and from 29.01° N to 37.77° N in latitude; Domain D02 (blue box) ranges from 118.93° E to 123.02° E in longitude and from 29.97° N to 35.97° N in latitude; Domain D03 (green box) ranges from 119.97° E to 122.01° E in longitude and from 31.50° N to 33.48° N in latitude. The yellow star indicates the location of wind farm.

For the physical parameterizations, the Kain-Fritsch (KF) cumulus scheme [44] was used to account for the unresolved convective effects, especially in the coarser parent domain, as it is a mature mass-flux convection scheme widely used in numerical weather prediction. This choice is physically motivated for mesoscale grids that cannot explicitly resolve deep convection fully and has been shown in previous studies to influence wind-speed simulations through the change in convection-related tendencies and pressure-gradient responses [45,46]. The WRF Single-Moment 6-Class (WSM6) scheme was used for cloud microphysics, which explicitly includes graupel and incorporates revised ice microphysics. This scheme provides a practical balance between computational efficiency and the representation of cloud, precipitation, and ice-cloud-radiation feedback processes [47]. The planetary boundary layer was represented by the TKE-based local closure scheme, the Mellor-Yamada-Nakanishi-Niino (MYNN) scheme [48], designed for WRF applications. This choice is particularly relevant to wind-energy prediction as hub-height wind forecasts are very sensitive to PBL treatment [49] and PBL parameterization changes can significantly change vertical mixing, momentum transport, wind shear, and turbulence structure [46,50].

These physics choices also matter for the accuracy of wind power forecasts. Errors introduced by alternative parameterization schemes in ambient wind speed, vertical wind shear, and turbulence characteristics can propagate nonlinearly to power prediction errors. Previous WRF-based wind-farm studies have shown that the skill of power prediction depends primarily on the model's ability to reproduce the ambient wind speed accurately [51]. Further sensitivity studies indicate that the sensitivity of wind simulation to physical parameterizations is often the most pronounced for PBL schemes, followed by microphysics. However, the impact of cumulus parameterization cannot be ignored for mesoscale simulation and can be more pronounced in the case of convective or extreme-weather states [45,46,49,50]. Critically, the optimal mix is not universal but is dependent on the region, season, terrain, and weather regime [46,50]. Thus, the present configuration should be considered as a physically motivated and internally consistent configuration

for this study, not as a globally best configuration for all wind-speed and wind power forecasting scenarios.

The longwave and shortwave radiation schemes use the Rapid Radiative Transfer Model (RRTM) scheme [52] and the Dudhia scheme [53], respectively. Other parameterization schemes include the RUC land surface scheme [54]. Except for the land surface scheme, which differs from that in Reference [4], all other parameterization settings are identical to those in that study.

3.2. Performance of Wind Speed Correction Model

To mitigate the wind speed bias in WRF forecasts, we construct a hybrid LSTM–LightGBM correction model. It utilizes a three-stage cascaded framework. Table 1 shows the comparison of the three metrics (R^2 , RMSE, MAE) between the WRF model and the correction model. The WRF model has considerable bias in wind speed forecasts, with an R^2 of only 0.5513. After the first LSTM correction module in the hybrid LSTM–LightGBM model, R^2 increases to 0.9401, while RMSE and MAE decrease to 0.7596 m/s and 0.5434 m/s, respectively, representing a substantial reduction in errors. By further incorporating LightGBM to correct the residuals, the model achieves an R^2 of 0.9798, while the RMSE and MAE fall to 0.4412 m/s and 0.2989 m/s. Finally, after SG smooth and physical constraints, the correction model achieves optimal performance, with an R^2 of 0.9867. Compared with the original WRF prediction, the hybrid correction model reduces RMSE by approximately 82.8% and MAE by about 84.8%, which demonstrates the effectiveness of the proposed multi-stage correction framework in wind speed forecast bias correction.

Table 1. The evaluation of the performance of the WRF and the hybrid LSTM–LightGBM model in the three stages.

	WRF	LSTM	LightGBM	Final
R^2	0.5513	0.9401	0.9798	0.9867
RMSE	2.0802	0.7596	0.4412	0.3575
MAE	1.5751	0.5434	0.2989	0.2394

The scatter plot of the fitting performance between the corrected results of the hybrid LSTM–LightGBM model and the observed values in the test set is visually shown in Figure 6a. The gray dots correspond to the raw WRF results, which show an obvious scattered distribution, especially in the moderate-to-high wind speed interval of 8–14 m/s. This phenomenon illustrates a large uncertainty related to the initial WRF predictions, which is also confirmed by a low R^2 value of 0.55 for the uncorrected forecasts. The red dots are much closer to the perfect prediction line, with R^2 for wind speed increasing to 0.9827. The data dispersion is significantly decreased after the LSTM–LightGBM model correction. The correction model successfully eliminates a considerable amount of random noise and significantly enhances prediction accuracy.

The wind speed bias distribution of the WRF model and correction model in the test set, shown in Figure 6b, further reveals the model’s capability in eliminating WRF systematic biases. The blue histogram and its fitted curve represent the baseline errors, showing a short and wide pattern, which implies a large standard deviation of errors and a wide distribution range, suggesting a high risk of prediction fluctuations. The red curve representing corrected errors features a pronounced, narrow peak precisely aligned with the zero-error line. This shows that the wind speed correction model keeps most errors within a small range of ± 1 m/s; both positive and negative mistakes in the original

forecasts are fixed. As a result, the average error is close to zero, making the predictions much more reliable.

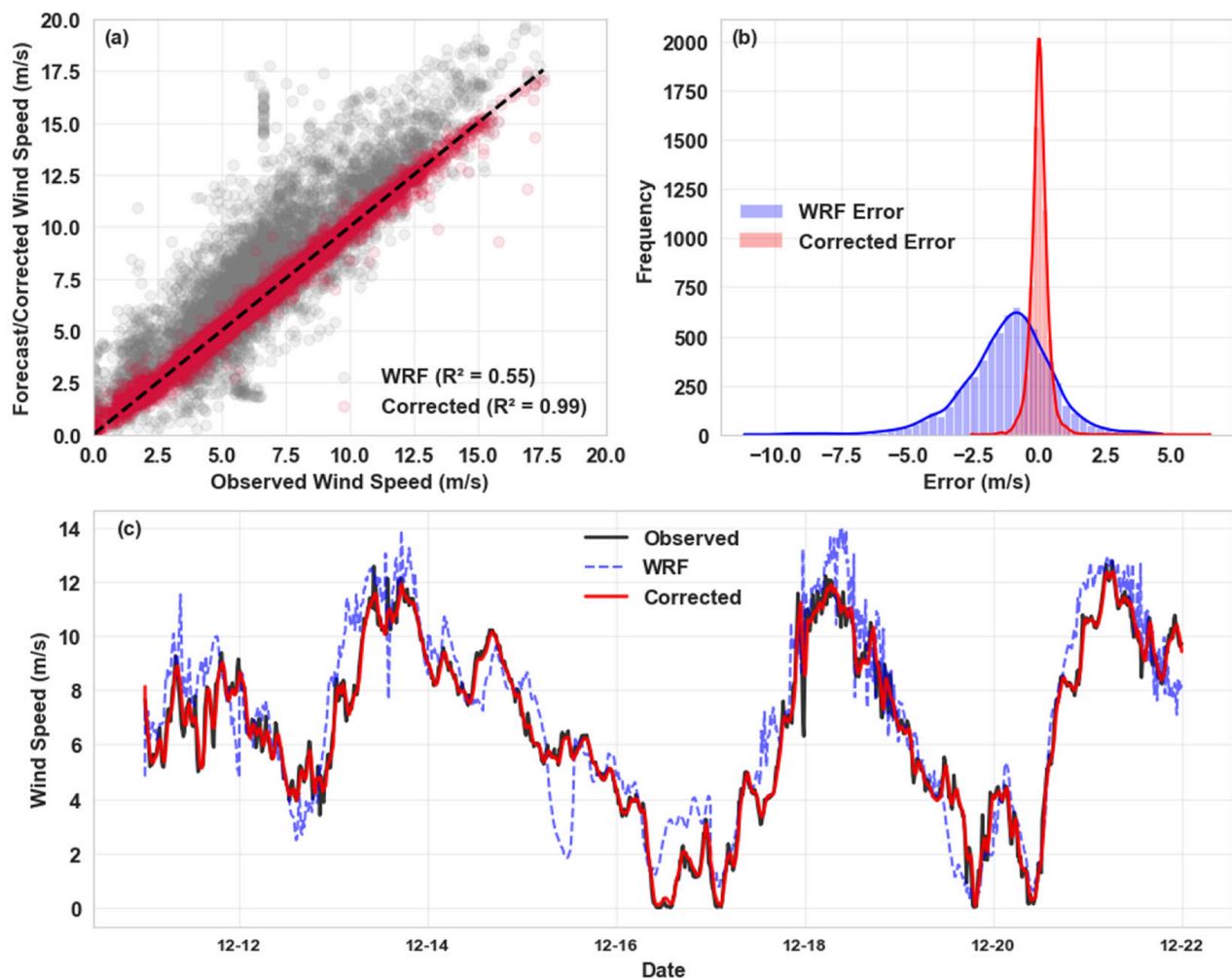


Figure 6. Performance of the hybrid LSTM–LightGBM correction model in wind speed forecasting. (a) A scatter plot illustrates the correlation between observed and predicted wind speeds in the test set, with WRF forecasts depicted in gray and the hybrid AI-corrected forecasts in red. (b) Histograms present the error distribution with WRF and the hybrid LSTM–LightGBM correction model in the test set. (c) A time series comparison from December 11 to 22, 2024, displaying observations (black), WRF forecasts (blue), and AI-corrected forecasts (red), along with the corresponding dates.

Figure 6c illustrates the variation in wind speed forecasted by the WRF and LSTM–LightGBM model during 12–22 December 2024. The black curve of actual observations undergoes drastic fluctuations. The blue WRF forecast curve, while showing the general trend, has noticeable phase shifts and differences in amplitude. In contrast, benefiting from the temporal feature memory capability of the LSTM module, the red corrected curve almost perfectly matches the peaks and troughs of the actual curve. While the MAE decreases from 1.1663 m/s to 0.2457 m/s, corresponding to a remarkable MAE reduction of 78.94%.

The bias represents the discrepancy between predictions and observations. A positive (negative) bias indicates overprediction (underprediction). Figure 7a presents the bias distribution of the WRF and the LSTM–LightGBM hybrid correction model across various wind speed ranges. The original WRF forecasts exhibit an overall positive bias, with the model overestimating wind speed by 0.82 to 1.49 m/s over all wind speed ranges. This overestimation is especially pronounced under low wind conditions (0–2 m/s), where the

bias reaches a maximum of 1.49 m/s. After correction by the LSTM–LightGBM model, the wind speed bias becomes more symmetric and much closer to zero. The correction model yields only a small positive bias of 0.25 m/s under low wind conditions (0–2 m/s), while the bias values are even smaller ($-0.04 \sim -0.01$ m/s) under medium wind conditions (4–12+ m/s). The mean bias of the model is reduced from 1.182 m/s to -0.007 m/s after correction, which corresponds to slight underestimation under most wind speed conditions. The developed model achieves a remarkable reduction in forecasting errors and effectively improves the original systematic bias pattern.

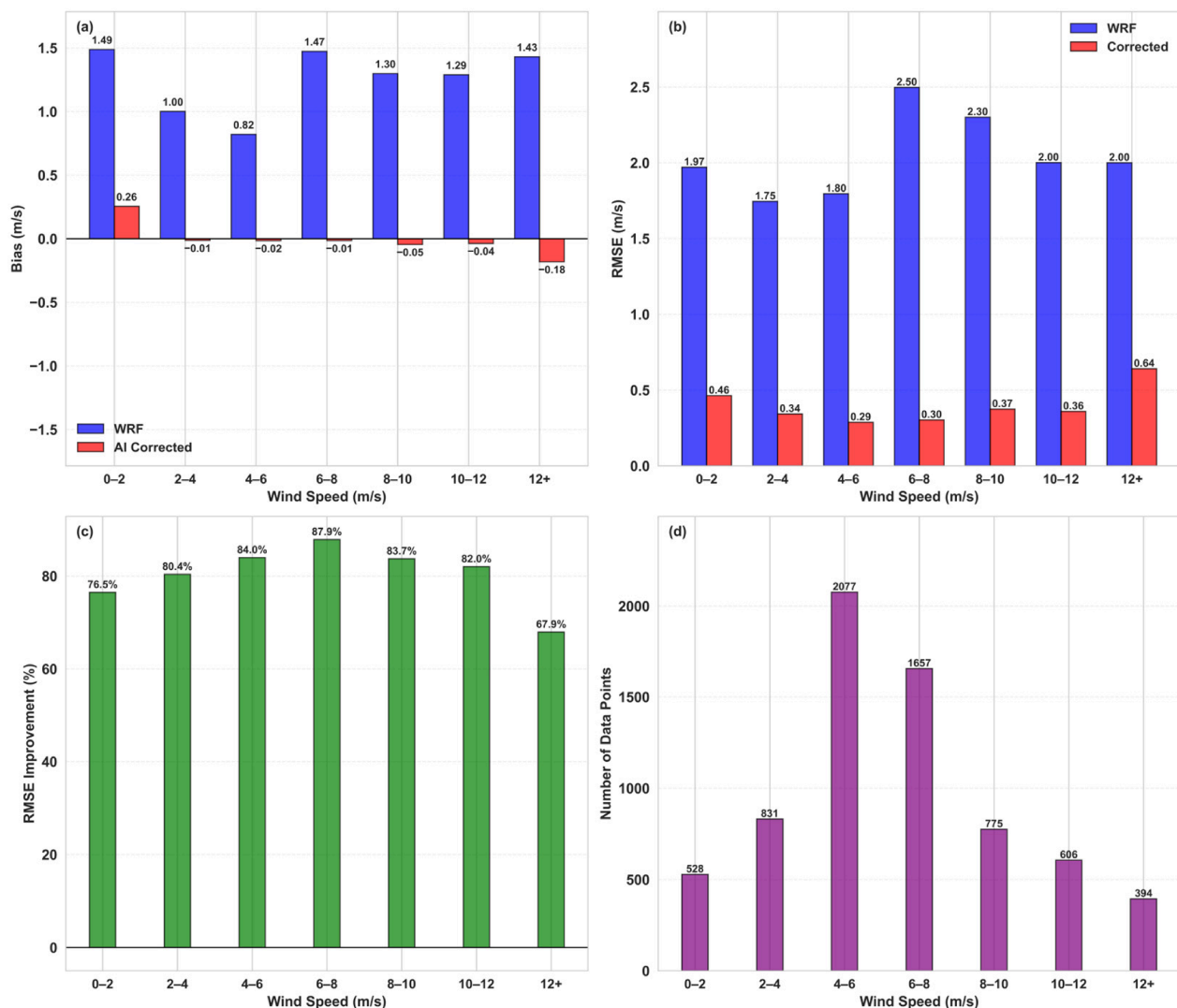


Figure 7. Wind speed bin analysis of forecast performance before and after correction. (a) Mean bias for WRF forecasts (blue) and AI-corrected forecasts (red) across different wind speed ranges. (b) RMSE comparison across different wind speed ranges. (c) Percentage reduction in RMSE achieved by the correction model. (d) Wind speed data distribution different wind speed ranges, showing the number of observations in each wind speed bin.

Figure 7b compares the wind speed RMSE of the WRF and the LSTM–LightGBM correction model within different wind speed ranges. For the original WRF forecasts, the wind speed RMSE ranges from 1.75 to 2.50 m/s, and the RMSEs are much larger under high wind conditions. After wind speed bias correction, the RMSEs are substantially reduced across all wind speed ranges, with RMSE values dropping to just 0.29–0.63 m/s. In the 6–8 m/s wind speed range, which accounts for 24.2% of the data, the RMSE decreases from

2.50 m/s to 0.30 m/s, corresponding to a relative reduction of 87.8%. The improvement in the hybrid LSTM–LightGBM model’s forecasting in the medium wind speed range is the most significant, and this range coincides with the most economically valuable operating conditions for wind farms.

Figure 7c further quantifies the performance improvement of the hybrid LSTM–LightGBM correction model using the relative RMSE reduction percentage. The model achieves positive improvements across all wind speed ranges (with reduction rates ranging from 68.4% to 87.8%). The green bars clearly reflect the extent of the model’s wind speed correction improvements, showing a distinct inverted U-shaped pattern. The RMSE reduction percentage is 76.9% under low wind conditions (0–2 m/s), gradually increases with rising wind speed, peaks at 87.8% in the 6–8 m/s range, and decreases to 68.4% under high wind conditions (12+ m/s). This distribution may be related to the following two factors: First, the training samples under medium wind conditions are the most abundant (accounting for 65.8% in total), which allows the correction model to learn more refined correction rules. Second, extreme wind events involve more complex physical mechanisms and limited sample representativeness (only 5.7%), which restricts the model’s correction capability.

Figure 7d shows the data distribution of the test set. The wind speed data are mainly distributed in the range of 4–10 m/s, accounting for 65.8% of the total samples. This range is consistent with the cut-in to rated wind speed range typically found in most wind farms, of which the 4–6 m/s range represents the largest proportion (30.3%). The above wind speed distribution is basically consistent with the optimal performance range of the correction model (4–10 m/s), indicating that the model can provide stable and reliable correction results under wind conditions with the highest frequency of wind energy utilization and the greatest economic value. By contrast, the high wind speed range (12+ m/s) accounts for only 5.7% because of the low occurrence frequency of extreme wind events. Such discrepancies further highlight the difficulty of accurately characterizing rapid localized wind variability, which is crucial for high-accuracy wind power forecasting.

3.3. The Performance of the Wind Power Model

Figure 8 shows the comparison between observed and predicted wind power by the new physics-guided and deep learning-integrated wind power forecasting model from October 21 to November 20, 2024. The overall variation in the model-predicted wind power agrees with the observation, indicating that the new wind power prediction model can reproduce the temporal variation in the wind power. Between October 24 and 26, the prediction curve shows several peaks significantly higher than the observed values, indicating a systematic overestimation tendency of the model during certain high wind power periods. In terms of evaluation metrics, the RMSE of the predicted wind power is 145.02 kW, the bias is 7.46 kW. The wind power forecasting model have achieved satisfactory overall prediction accuracy and can effectively characterize the power variation trends. However, there is still room for further optimization, particularly during periods of rapid power fluctuations.

The goodness of fit of the predicted wind power to the actual values over the test set is also illustrated in Figure 9a. The majority of data points converge closely to the red dashed ideal prediction line, which verifies good agreement between actual and predicted wind power data. The green solid line is the linear regression fit, and its slope indicates the systematic bias trend of overestimation or underestimation of the model. On. For low and medium wind power conditions (0–2000 kW), data samples are densely distributed along the diagonal line, manifesting high prediction accuracy under conventional operating scenarios. In contrast, some data points are significantly deviating from the ideal prediction line in the high-power range (above 3000 kW). When actual wind power is relatively

high, the predicted values are mostly distributed above the ideal line, revealing slight overestimation by the wind power forecasting model during high-power periods. The correction model achieves an MAE of 103.94 kW, with a small average prediction bias. The R^2 reaches as high as 0.9853, indicating that the wind power forecasting model can explain 98.53% of the actual power variation, with a strong linear correlation between the predictions and measurements. The negative mean bias of -7.21 kW further confirms the existence of a slight systematic underestimation in this model.

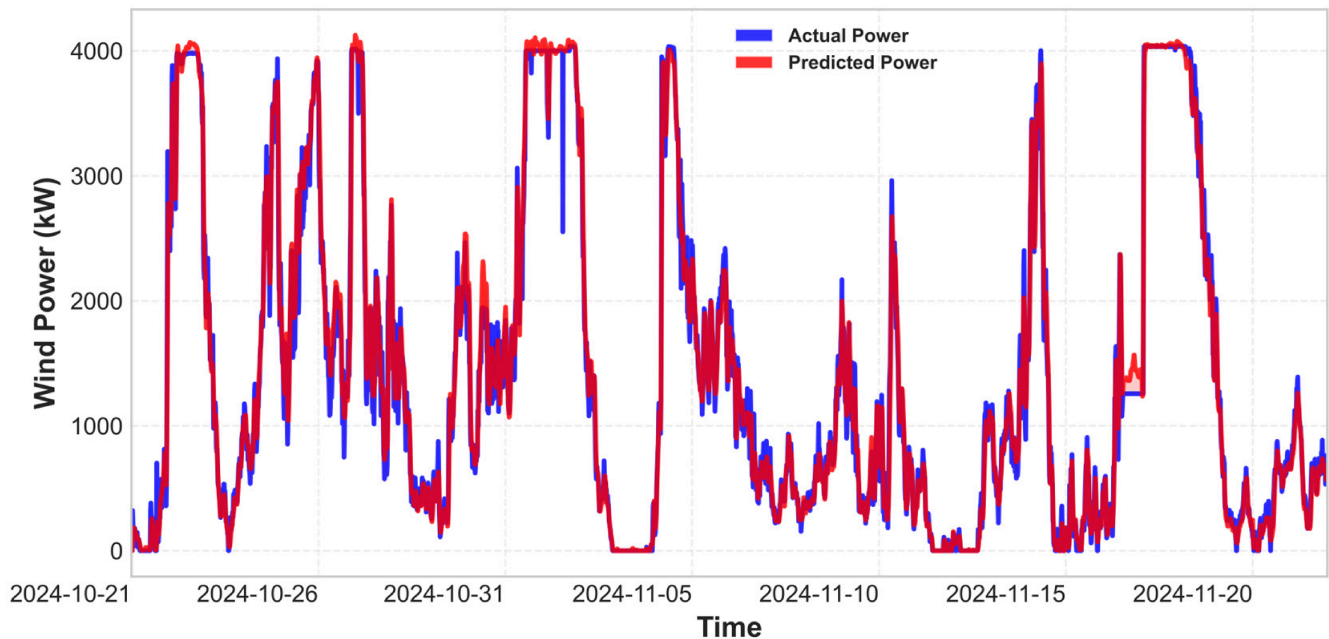


Figure 8. Time series comparison of actual versus predicted power from October 21 to November 20, 2024. The blue line represents actual power, and the red line represents predicted power.

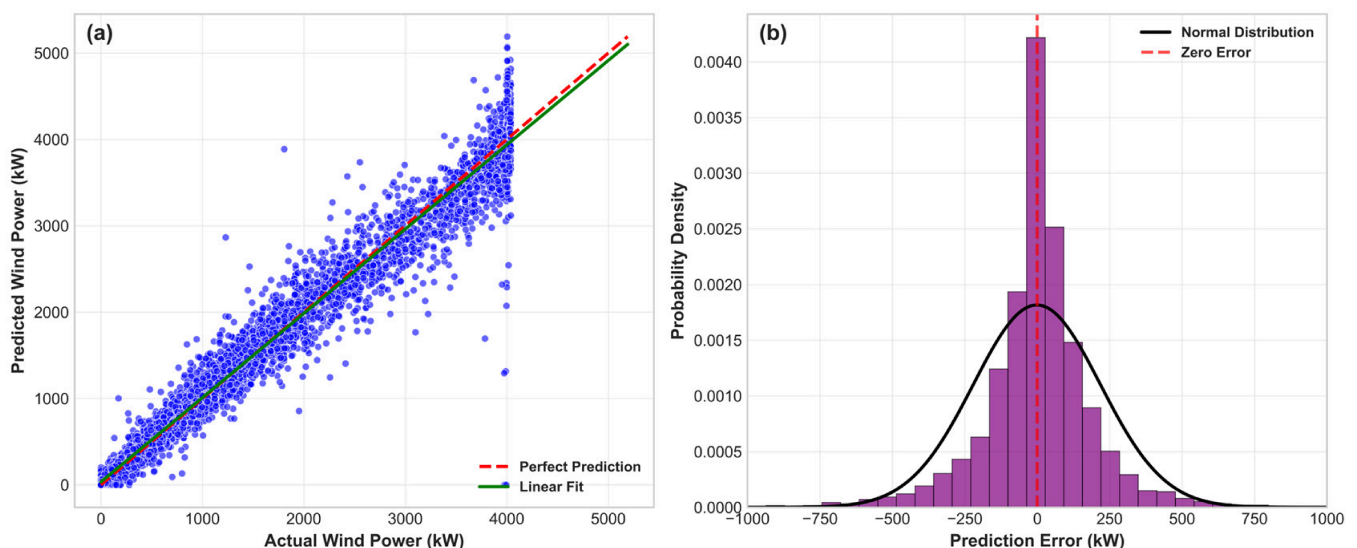


Figure 9. Performance assessment of the wind power prediction model in the test set. (a) The scatter plot compares the predicted and observed wind power values, along with a perfect prediction line (red dashed) and a linear regression fit (green solid). (b) The error distribution histogram depicts the frequency of prediction errors, overlaid with a fitted normal distribution curve (black solid) to aid in assessing the normality of the error distribution.

Figure 9b further illustrates the statistical characteristics of the forecasted wind power errors for the proposed wind power forecasting model. The error histogram is mainly distributed around zero, which indicates the good performance of the model on most samples. However, the distribution is slightly shifted towards negative values, consistent with the slight underestimation tendency revealed by Figure 9a. The negative skewness of -0.4534 suggests a moderately longer left tail, indicating that underestimation errors are somewhat more prevalent than overestimation errors. In addition, the kurtosis is 14.9822 , which is much higher than the theoretical value of 3 for a standard normal distribution [55]. This result means that the distribution of the forecast error is highly leptokurtic with a sharp peak at the center and fat tails. This implies that most of the prediction errors are small, but some samples with large errors are present. Thus, the forecast errors are not strictly normally distributed, and the model errors are negatively skewed and have significant heavy-tail characteristics.

4. Discussion

Although metrics such as RMSE and MAE can assess the prediction capability of the wind power model from a macroscopic perspective, they can hardly explain the intrinsic reasoning behind specific predictions or determine the key features for the model. The interpretability deficiency of most artificial intelligence models weakens the confidence of users in the predictive results and restricts the application of the corresponding decision strategies. Interpretability drawbacks of artificial intelligence models need to be addressed for the safety and transparency of power system decision-making [56]. These limitations are overcome in the following sections with permutation feature importance and physical coefficient uncertainty analysis. These methods can help to clarify the internal mechanism of the model, to quantify the contribution of constructed features, and to verify the stability of the three physical coefficients during the training process of the model.

Permutation feature importance [57] is a popular choice because it is computationally inexpensive, can be used with many types of models, and is easy to interpret. The idea is simple: shuffle the values of one feature randomly and observe how the prediction performance of the model drops. If the feature is important, the error will change a lot; if it is irrelevant, shuffling will not change anything. Typically, each feature is randomly permuted several times, and the importance score is averaged over these repetitions to get a stable and reliable result.

As shown in Table 1, the introduction of LightGBM raises the R^2 from 0.9401 to 0.9798. Since additional features are incorporated into this module, feature importance analysis can identify which specific features drive this improvement. Top-10 feature importance from the LightGBM residual-correction model is shown in Figure 10. According to the feature importance ranking, temporal autocorrelation features play a dominant role. The residual at lag 1 (residual_lag1) ranks highest among all features. The correction error at the previous time step has a decisive impact on the current error. The subsequent features lstm_diff1 (first-order difference in the LSTM-corrected wind speed) and u_fc_diff1 (first-order difference in the WRF raw forecast) also show significantly higher importance than other features, suggesting that the temporal trend terms of wind speed underpin model correction. Features including residual_rolling_mean_4 (4-step rolling mean of residuals) and residual_lag2 further demonstrate that local temporal patterns over short time windows play a key role in predicting model residuals. Meanwhile, lstm_adjustment (the difference between the raw forecast and the LSTM-corrected wind speed) and wind_tc_ratio (the ratio of LSTM-corrected wind speed to absolute air temperature) both appear in the top 10 most important features. Notably, the primitive meteorological predictors—air temperature, humidity, and atmospheric pressure—are excluded from the top 10 most influential features.

This indicates that their inherent predictive information has been fully and effectively captured by the LSTM model through the constructed lag and difference features.

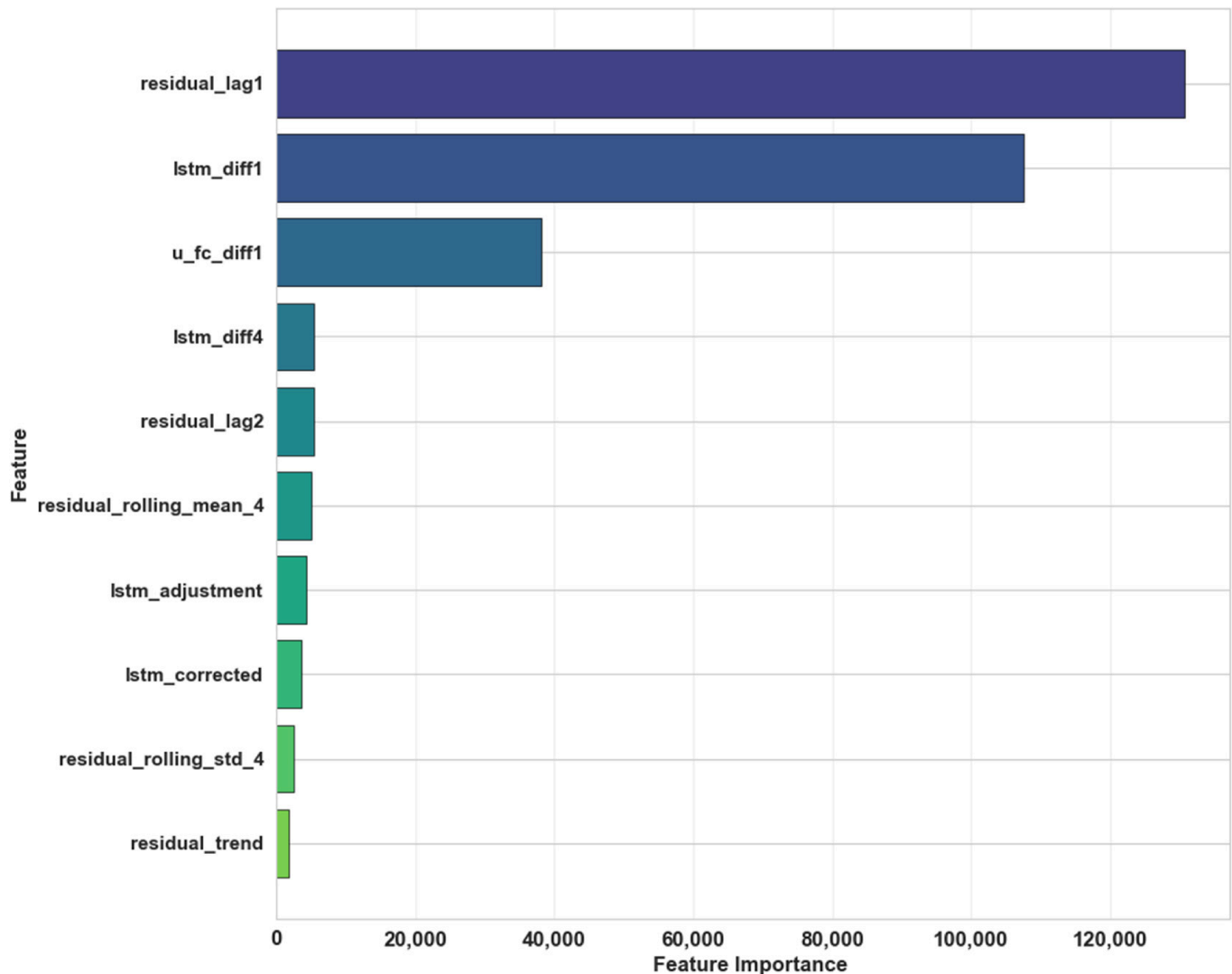


Figure 10. Top-10 feature importance (based on gain) from the LightGBM residual-correction model.

Figure 11 displays the average increase in MAE when each feature is randomly permuted, which quantifies the key features in predicting the wind power. Both wind speed and density power terms are leading predictive factors. The permutation of wind speed leads to an MAE increase of 0.9141, and the shuffling of the density power term results in an MAE increase of 0.7597. Power ramp and acceleration contribute moderately with an increase in MAE of 0.5555 and 0.2578, respectively, and are ranked second and third in the importance of the features, outperforming commonly used meteorological parameters such as temperature, humidity, and pressure. In addition to wind speed, physical factors that characterize power variation patterns and inertial response are also crucial to ensure accurate predictions. Meanwhile, time-encoding features (Time Sin/Cos) show marginal importance, suggesting that the model can independently capture periodic temporal patterns via physical feature learning without additional temporal encoding.

To further explore the model's internal learning properties, Figure 12 illustrates the parametric stability and behavioral variability across different random initializations. All three constructed physical coefficients yield positive values that are highly consistent with the mean predictions, demonstrating that the embedded physics correction layer enables

the model to precisely perceive and replicate wind energy conversion mechanisms. The density power coefficient dominates the three indicators with a mean value of 0.7295, acting as the primary driver of wind power fluctuations. The wind acceleration coefficient with a mean of 0.5313 quantifies how changes in wind speed influence wind power. Positive wind speed acceleration can stimulate the inertial response of turbine blades and generate extra power fluctuations, and the model effectively learns this positive gust-driven effect. Moreover, the mean power ramp rate of 0.5102 effectively reflects the control lag in the pitch and yaw systems under complex wind conditions. Notably, the standard deviations of all three coefficients stabilize at the order of 10^{-2} , demonstrating the high stability of the model training process and effectively eliminating common training defects such as coefficient vanishing and sign reversal.

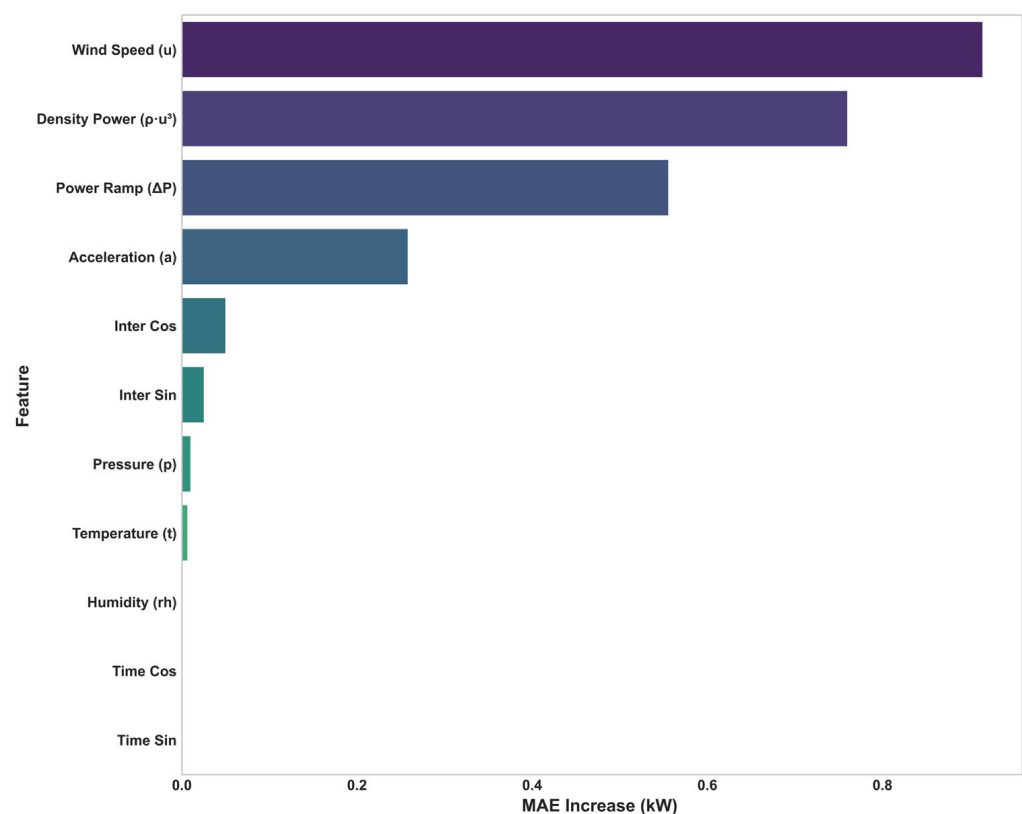


Figure 11. Feature importance analysis via permutation impact. The bar chart displays the mean increase in MAE when each feature is randomly permuted, quantifying feature importance for wind power prediction.

An ablation study was performed to assess the contribution of each main component within the proposed framework, building the model up from the baseline by incrementally adding different modules. Specifically, four variants were evaluated: GRU, GRU + CNN, GRU + CNN + Transformer, and GRU + CNN + Transformer + Physics Correction. The worst performance was obtained by the model based on the GRU only, as shown in Table 2, indicating that the recurrent structure alone cannot learn the complex nonlinear characteristics of wind power variation. The introduction of the CNN module greatly improved the prediction accuracy, and the RMSE and MAE were reduced to 170.49 kW and 106.80 kW, respectively, and the R^2 increased to 0.9846. The result shows that extracting the local temporal pattern plays an important role in improving the forecasting performance. The integration of the Transformer module is further applied and yields additional gains by reducing RMSE to 167.69 kW and MAE to 105.26 kW, and increasing R^2 to 0.9851, indicating that the modeling of long-range temporal dependencies is complementary

to convolutional feature learning. The full model with Physics Correction performed best overall with an RMSE of 166.70 kW, an MAE of 103.94 kW, and an R^2 of 0.9853. The improvement from the physics correction branch was rather small compared to the CNN module. However, it consistently improved the predictive results, showing that incorporating physically guided correction into the data-driven framework can further improve model reliability and accuracy. Overall, these results confirm the effectiveness of the proposed architecture and prove that each introduced module contributes positively to the final forecasting performance.

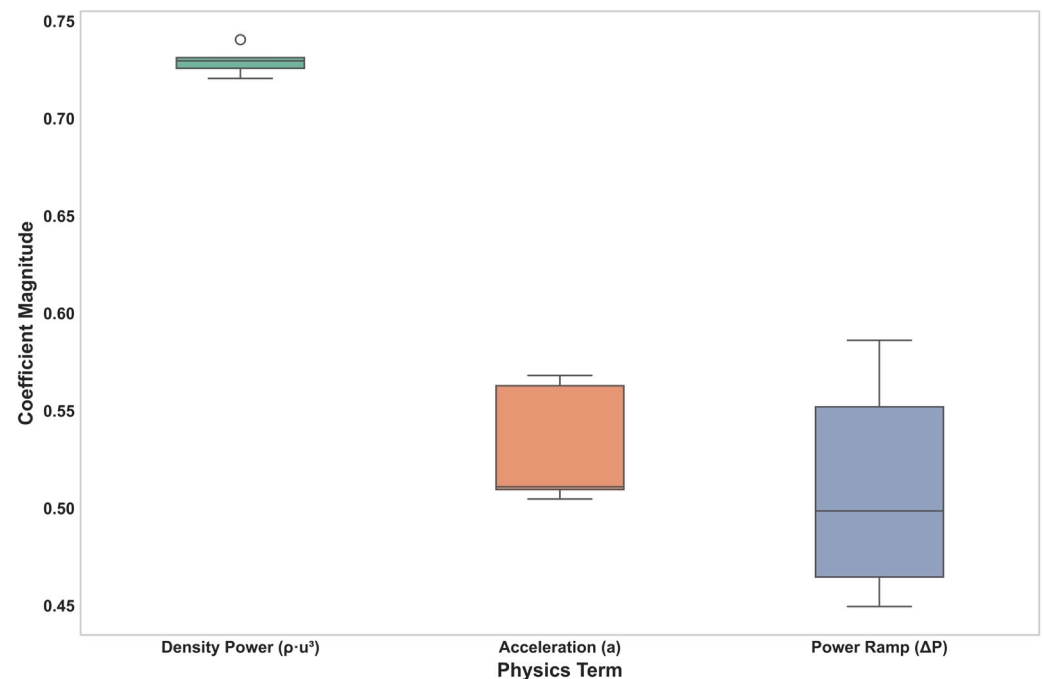


Figure 12. Distribution of three learned physics coefficients for five training Seeds, boxplot showing the variability of the physics coefficients from five independent model initializations.

Table 2. Ablation study results of different model variants.

Model	RMSE	MAE	R^2
GRU	1195.62	910.00	0.2419
+CNN	170.49	106.80	0.9846
+Transformer	167.69	105.26	0.9851
+Physics Correction	166.70	103.94	0.9853

To verify the effectiveness of the proposed model, we compared it with several representative baseline methods widely used in wind power prediction. We chose XGBoost as a strong machine-learning baseline due to its ability to perform nonlinear regression. LSTM and BiLSTM were used as traditional recurrent architectures for modeling temporal dependencies. CNN + LSTM was included as a hybrid baseline to learn local patterns and sequential dynamics. A transformer was used as an advanced attention-based model that can learn long-range temporal dependencies. For a fair comparison, all models were trained and tested on the same dataset using the same experimental settings. The predictive performance of the models was evaluated in terms of RMSE, MAE, and R^2 . As shown in Table 3, the proposed physics-guided model achieved the best overall performance with an RMSE of 166.70 kW, MAE of 103.94 kW, and R^2 of 0.9853. Among the baseline models, Transformer and CNN + LSTM presented relatively competitive results with RMSE values of 170.31 kW and 170.72 kW, respectively, showing that both attention-based mechanisms

and hybrid feature extraction strategies are effective for wind power forecasting. XGBoost also showed good performance with an RMSE of 171.26 kW and an MAE of 104.96 kW, demonstrating that nonlinear machine-learning methods are still competitive for this problem. On the other hand, the results of LSTM and BiLSTM are worse, especially the LSTM alone, which has an RMSE of 197.31 kW and an R^2 of 0.9794, showing that the recurrent architectures alone are not enough to describe the complex variability of wind power. In general, the improvements of the proposed model on the best deep-learning baselines were moderate but consistent across all evaluation metrics, showing the proposed framework offers a more accurate and robust solution to wind power forecasting.

Table 3. Comparison of the proposed physics-guided model with representative baseline methods for wind power forecasting.

Model	RMSE	MAE	R^2
XGBoost	171.26	104.96	0.9844
LSTM	197.31	127.60	0.9794
BiLSTM	179.98	116.20	0.9828
CNN + LSTM	170.72	105.75	0.9845
Transformer	170.31	109.99	0.9846
Physics-guided	166.70	103.94	0.9853

5. Conclusions

A wind power prediction model system integrating the WRF numerical weather model and artificial intelligence technology is constructed to achieve high-precision wind power forecasting. The WRF model is employed first to obtain wind speed forecasts with high spatiotemporal resolution. Subsequently, a hybrid LSTM–LightGBM model is developed to correct the WRF-derived wind speed, which effectively reduces the forecast bias of the WRF model and supplies high-precision wind speed for subsequent wind power prediction. During the test period, the RMSE of the WRF wind speed forecasts is reduced by up to 82.8%, and the R^2 is improved from 0.55 to 0.99. The large number of training samples in the medium wind speed range (accounting for 65.8% cumulatively) enables the correction model to learn more refined correction rules. As a result, the wind speed correction model yields the most significant forecast improvement in the medium wind speed range, which exactly corresponds to the most economically valuable operating conditions of wind farms.

After correcting the wind speed bias in the WRF model, this study further develops a physics-guided and deep learning-integrated model. By introducing physical features such as “density power”, wind power ramp rate and wind acceleration, the model’s generalization performance and interpretability are greatly improved. The model demonstrates high forecasting accuracy, with an MAE of 103.94 kW and an R^2 of 0.9853 for wind power. This study also reveals the physical learning mechanism of the deep learning model in wind power prediction from the perspective of interpretability through permutation feature importance and physical coefficient uncertainty analysis methods. It confirms that the model not only effectively captures the cubic law of wind power versus wind speed but also exhibits physically consistent learning behaviors under different training initializations. This study provides viable technical references for solving practical problems concerning prediction accuracy, operational reliability, and model interpretability in meteorological-energy cross-domain wind power forecasting tasks.

The proposed wind power forecasting model exhibits three prominent advantages over conventional models. Firstly, the integration of physical laws and data-driven features alleviates the black-box problem inherent in pure data-driven models and enables prediction results to better conform to engineering practices. Secondly, the synergistic use of

ensemble learning and composite loss functions significantly enhances the model's robustness and stability. In addition, permutation importance analysis and physical coefficient uncertainty evaluation are introduced to quantify feature contributions to wind power prediction and characterize the stability and variability of physical knowledge learning during repeated model training. This establishes an observable physical diagnosis mechanism for conventional black-box models.

The effectiveness of the proposed framework is further validated by the comparative experiments and ablation analyses. Compared with representative baseline models, including XGBoost, LSTM, BiLSTM, CNN + LSTM, and Transformer, the proposed physics-guided model achieved the best overall performance, demonstrating its superiority in wind power forecasting. Further ablation study shows that each module contributes positively to the final prediction performance. The CNN branch makes the largest contribution to the improvement by focusing on better local pattern extraction, and the Transformer branch further improves the modeling of long-term temporal dependencies. The physics-guided correction branch achieves further consistent gains, indicating that the introduced physical features enhance the accuracy and robustness of the model. The results obtained corroborate that the proposed model not only achieves better performance compared to conventional benchmark methods but also effectively integrates multiple functional components.

The presented wind power forecasting model overall has a certain level of prediction accuracy and can effectively reflect the trend of wind power changes. However, during certain high wind power periods, the model tends to systematically overestimate, indicating that there is still room for optimization, especially during rapid power change phases. To improve the model's robustness and prediction accuracy under extreme conditions, our future research will aim to improve the prediction algorithm for high-power intervals and conduct a more detailed analysis of typical operating conditions or weather patterns that result in significant overestimation. Consequently, these improvements will enhance the model's robustness and prediction accuracy under extreme conditions. To address the larger errors under high wind speeds (>12 m/s) due to sample scarcity, we will propose adding small Gaussian noise (1–3%) and scaling perturbations ($\pm 5\%$) to existing high-wind samples, bounded by cut-in/out speeds, to generate realistic augmented data and improve generalization in future work.

In this study, only the complete observation data set in 2024 was used for model training and validation. A one-year dataset is not considered sufficient proof for full interannual generalizability of the proposed model, but is considered sufficient for an initial site-specific validation. This is because the 2024 data set covers the full annual cycle of local wind conditions, including seasonal variability and representative strong wind events, and therefore allows for a systematic evaluation of model performance under the wind regime observed at this site during this year. This treatment is consistent with common practice in the wind power forecasting literature, where models have been developed and validated using often one-year datasets or even shorter records. For example, Dong et al. put forward a regional wind power forecasting method for small data sets and achieved reliable 24 h forecasting performance with only three months of NWP data [58]. Lin et al. directly compared one-month and one-year training data sets for wind power forecasting [59]. Aftab et al. proposed a wind power forecasting model with one-year of SCADA and NWP data [60]. Karaman also developed and tested wind power prediction models on a full year data-set of the year 2018 [61]. These studies indicate that a single-year dataset can support an initial validation of methodological effectiveness in wind power forecasting, although the resulting conclusions are generally limited to the specific period and scenario represented by the available data.

However, the absence of multi-year continuous samples prevents a complete evaluation of the interannual robustness of the proposed correction model. Previous work has shown that distribution shift in time-varying wind power data can impede model generalization to future periods [62]. Thus, the results of the present study should be interpreted as showing the feasibility and effectiveness of the proposed method under the 2024 wind regime at this site, rather than proving its universal robustness under the wind conditions in other years. In future work, longer continuous observation records will be collected at this site for out-of-year validation, robustness assessment under different annual wind regimes, and further evaluation of the broader engineering applicability of the proposed model.

Author Contributions: Conceptualization, B.W., M.Z. (Mingjian Zeng) and H.S.; methodology, B.W., M.Z. (Ming Zhang); software, Q.C.; validation, M.Z. (Ming Zhang); formal analysis, B.W.; investigation, B.W.; resources, B.W.; data curation, M.Z. (Ming Zhang); writing—original draft preparation, B.W.; writing—review and editing, M.Z. (Mingjian Zeng) and H.S.; visualization, Q.C.; supervision, M.Z. (Mingjian Zeng) and H.S.; project administration, M.Z. (Mingjian Zeng) and H.S.; funding acquisition, M.Z. (Mingjian Zeng) and H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (No. 2023YFC3008201), the National Natural Science Foundation of China (42376026), the scientific research project of China Meteorological Administration (CMAJBGS202315), and the scientific research project of Jiangsu Meteorological Bureau (KM202503).

Data Availability Statement: The data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chen, Y.; Zhang, S.; Zhang, W.; Peng, J.; Cai, Y. Multifactor spatio-temporal correlation model based on a combination of convolutional neural network and long short-term memory neural network for wind speed forecasting. *Energy Convers. Manag.* **2019**, *185*, 783–799. [\[CrossRef\]](#)
2. Ma, Y.; Yu, L.; Zhang, G. Short-term wind power forecasting with an intermittency-trait-driven methodology. *Renew. Energy* **2022**, *198*, 872–883. [\[CrossRef\]](#)
3. Yang, M.; Che, R.; Yu, X.; Su, X. Dual NWP wind speed correction based on trend fusion and fluctuation clustering and its application in short-term wind power prediction. *Energy* **2024**, *302*, 131802. [\[CrossRef\]](#)
4. Prósper, M.A.; Otero-Casal, C.; Fernández, F.C.; Miguez-Macho, G. Wind power forecasting for a real onshore wind farm on complex terrain using WRF high resolution simulations. *Renew. Energy* **2019**, *135*, 674–686. [\[CrossRef\]](#)
5. Elshafei, B.; Popov, A.; Giddings, D. Enhanced offshore wind resource assessment using hybrid data fusion and numerical models. *Energy* **2024**, *310*, 133208. [\[CrossRef\]](#)
6. Tuncar, E.A.; Sağlam, Ş.; Oral, B. A review of short-term wind power generation forecasting methods in recent technological trends. *Energy Rep.* **2024**, *12*, 197–209. [\[CrossRef\]](#)
7. Xie, C.; Yang, X.; Chen, T.; Fang, Q.; Wang, J.; Shen, Y. Short-term wind power prediction framework using numerical weather predictions and residual convolutional long short-term memory attention network. *Eng. Appl. Artif. Intell.* **2024**, *133*, 108543. [\[CrossRef\]](#)
8. Couto, A.; Estanqueiro, A. Enhancing wind power forecast accuracy using the weather research and forecasting numerical model-based features and artificial neuronal networks. *Renew. Energy* **2022**, *201*, 1076–1085. [\[CrossRef\]](#)
9. Skamarock, W.C.; Klemp, J.B.; Dudhia, J.; Gill, D.O.; Liu, Z.; Berner, J.; Wang, W.; Powers, J.G.; Duda, M.G.; Barker, D.M. *A Description of the Advanced Research WRF*, 4th ed.; National Center for Atmospheric Research: Boulder, CO, USA, 2019; p. 145.
10. Brotzge, J.A.; Berchoff, D.; Carlis, D.L.; Carr, F.H.; Carr, R.H.; Gerth, J.J.; Gross, B.D.; Hamill, T.M.; Haupt, S.E.; Jacobs, N.; et al. Challenges and opportunities in numerical weather prediction. *Bull. Am. Meteorol. Soc.* **2023**, *104*, E698–E705. [\[CrossRef\]](#)
11. Zaman, T.; Juliano, T.W.; Kosovic, B.; Astitha, M. Assessment of offshore wind and turbulence prediction for the northeast US using a three-dimensional PBL parameterization in the WRF model at gray zone resolution. *J. Geophys. Res. Atmos.* **2026**, *131*, e2025JD044765. [\[CrossRef\]](#)

12. Di Paola, F.; Cimini, D.; De Natale, M.P.; Gallucci, D.; Gentile, S.; Gerdali, E.; Larosa, S.; Nilo, S.T.; Ricciardelli, E.; Ripepi, E.; et al. Wind Speed Downscaling of the WRF Model at Subkilometer Scale in Complex Terrain for Wind Power Applications. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 9139–9177. [[CrossRef](#)]
13. Mo, J.; Shen, Y.; Yuan, B.; Li, M.; Ding, C.; Jia, B.; Ye, D.; Wang, D. Assessment of Numerical Forecasts for Hub-Height Wind Resource Parameters during an Episode of Significant Wind Speed Fluctuations. *Atmosphere* **2024**, *15*, 1112. [[CrossRef](#)]
14. Xu, J.; Xiao, Z.; Lin, Z.; Li, M. System bias correction of short-term hub-height wind forecasts using the Kalman filter. *Prot. Control Mod. Power Syst.* **2021**, *6*, 37. [[CrossRef](#)]
15. Zhang, W.; Tian, M.; Hai, S.; Wang, F.; An, X.; Li, W.; Li, X.; Sheng, L. Improving the Forecasts of Coastal Wind Speeds in Tianjin, China Based on the WRF Model with Machine Learning Algorithms. *J. Meteorol. Res.* **2024**, *38*, 570–585. [[CrossRef](#)]
16. Zhao, R.; Dai, K.; Cao, Y.; Zhu, Y.; Wang, Y.; Wang, Q.; Xu, Y. Improved Short-Term Surface Wind Speed Forecasts Using Random Forest Approach with Sliding Time Window and Region Regression. *Wea Forecast.* **2025**, *40*, 1917–1934. [[CrossRef](#)]
17. Xiong, X.; Guo, X.; Zeng, P.; Zou, R.; Wang, X. A short-term wind power forecast method via XGBoost hyper-parameters optimization. *Front. Energy Res.* **2022**, *10*, 905155. [[CrossRef](#)]
18. Zeng, X.; Abdullah, N.; Liang, B. A widely applicable and robust LightGBM-Artificial neural network forecasting model for short-term wind power density. *Heliyon* **2023**, *9*, e23071. [[CrossRef](#)] [[PubMed](#)]
19. Duan, J.; Zuo, H.; Bai, Y.; Duan, J.; Chang, M.; Chen, B. Short-term wind speed forecasting using recurrent neural networks with error correction. *Energy* **2021**, *217*, 119397. [[CrossRef](#)]
20. Liu, X.; Lin, Z.; Feng, Z. Short-term offshore wind speed forecast by seasonal ARIMA—A comparison against GRU and LSTM. *Energy* **2021**, *227*, 120492. [[CrossRef](#)]
21. Memarzadeh, G.; Keynia, F. A new short-term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets. *Energy Convers. Manag.* **2020**, *213*, 112824. [[CrossRef](#)]
22. Niu, Z.; Yu, Z.; Tang, W.; Wu, Q.; Reformat, M. Wind power forecasting using attention-based gated recurrent unit network. *Energy* **2020**, *196*, 117081. [[CrossRef](#)]
23. Liang, Y.; Zhang, D.; Zhang, J.; Hu, G. A State-of-the-Art Analysis on Decomposition Method for Short-Term Wind Speed Forecasting Using LSTM and a Novel Hybrid Deep Learning Model. *Energy* **2024**, *313*, 133826. [[CrossRef](#)]
24. Yan, Y.; Wang, X.; Ren, F.; Shao, Z.; Tian, C. Wind speed prediction using a hybrid model of EEMD and LSTM considering seasonal features. *Energy Rep.* **2022**, *8*, 8965–8980. [[CrossRef](#)]
25. Li, Q.; Wang, G.; Wu, X.; Gao, Z.; Dan, B. Arctic short-term wind speed forecasting based on CNN-LSTM model with CEEMDAN. *Energy* **2024**, *299*, 131448. [[CrossRef](#)]
26. Zhang, Y.; Chen, B.; Pan, G.; Zhao, Y. A novel hybrid model based on VMD-WT and PCA-BP-RBF neural network for short-term wind speed forecasting. *Energy Convers. Manag.* **2019**, *195*, 180–197. [[CrossRef](#)]
27. Li, H.; He, S.; Yuan, J.; Wang, C. A Wind Power Prediction Method Integrating Dynamic Multi-Scale Spatio-Temporal Modelling, Adaptive Multi-Strategy Local Decomposition, and Meta-Learning Ensemble Model. *Energy* **2025**, *340*, 139329. [[CrossRef](#)]
28. Ateş, K.T. Estimation of Short-Term Power of Wind Turbines Using Artificial Neural Network (ANN) and Swarm Intelligence. *Sustainability* **2023**, *15*, 13572. [[CrossRef](#)]
29. Yuan, D.-D.; Li, M.; Li, H.-Y.; Lin, C.-J.; Ji, B.-X. Wind Power Prediction Method: Support Vector Regression Optimized by Improved Jellyfish Search Algorithm. *Energies* **2022**, *15*, 6404. [[CrossRef](#)]
30. Malakouti, S.M.; Ghiasi, A.R.; Ghavifekr, A.A.; Emami, P. Predicting wind power generation using machine learning and CNN-LSTM approaches. *Wind Eng.* **2022**, *46*, 1853–1869. [[CrossRef](#)]
31. Yu, M.; Zhang, Z.; Li, X.; Yu, J.; Gao, J.; Liu, Z.; You, B.; Zheng, X.; Yu, R. Superposition graph neural network for offshore wind power prediction. *Future Gener. Comput. Syst.* **2020**, *113*, 145–157. [[CrossRef](#)]
32. Sun, S.; Liu, Y.; Li, Q.; Wang, T.; Chu, F. Short-term multi-step wind power forecasting based on spatio-temporal correlations and transformer neural networks. *Energy Convers. Manag.* **2023**, *283*, 116916. [[CrossRef](#)]
33. Yang, M.; Huang, Y.; Xu, C.; Liu, C.; Dai, B. Review of Several Key Processes in Wind Power Forecasting: Mathematical Formulations, Scientific Problems, and Logical Relations. *Appl. Energy* **2025**, *377*, 124631. [[CrossRef](#)]
34. Kirchner-Bossi, N.; Kathari, G.; Porté-Agel, F. A Hybrid Physics-Based and Data-Driven Model for Intra-Day and Day-Ahead Wind Power Forecasting Considering a Drastically Expanded Predictor Search Space. *Appl. Energy* **2024**, *367*, 123375. [[CrossRef](#)]
35. Li, Z.; Xu, R.; Luo, X.; Cao, X.; Sun, H. Short-term wind power prediction based on modal reconstruction and CNN-BiLSTM. *Energy Rep.* **2023**, *9*, 6449–6460. [[CrossRef](#)]
36. Yang, Q.; Chen, K.; Man, J.; Duan, J.; Jin, Z. A Method for Cleaning Wind Power Anomaly Data by Combining Image Processing with Community Detection Algorithms. *Glob. Energy Interconnect.* **2024**, *7*, 293–312. [[CrossRef](#)]
37. Liu, S.; Zhang, Y.; Du, X.; Xu, T.; Wu, J. Short-Term Power Prediction of Wind Turbine Applying Machine Learning and Digital Filter. *Appl. Sci.* **2023**, *13*, 1751. [[CrossRef](#)]
38. Biswas, S.; Chen, J.S.J. Power Coefficient for Large Wind Turbines Considering Wind Gradient Along Height. *Energies* **2025**, *18*, 740. [[CrossRef](#)]

39. Castillo, O.C.; Andrade, V.R.; Rivas, J.J.R.; González, R.O. Comparison of Power Coefficients in Wind Turbines Considering the Tip Speed Ratio and Blade Pitch Angle. *Energies* **2023**, *16*, 2774. [\[CrossRef\]](#)
40. Li, W.; Jia, X.; Li, X.; Wang, Y.; Lee, J. A Markov model for short term wind speed prediction by integrating the wind acceleration information. *Renew. Energy* **2021**, *164*, 242–253. [\[CrossRef\]](#)
41. Sutarna, N.; Tjahyadi, C.; Oktivasari, P.; Dwiyani, M.; Tohazen, T. Hyperparameter Tuning Impact on Deep Learning Bi-LSTM for Photovoltaic Power Forecasting. *J. Robot. Control* **2024**, *5*, 677–693.
42. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-excitation networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 2011–2023. [\[CrossRef\]](#) [\[PubMed\]](#)
43. de Amorim, L.B.V.; Cavalcanti, G.D.C.; Cruz, R.M.O. The choice of scaling technique matters for classification performance. *Appl. Soft Comput.* **2023**, *133*, 109924. [\[CrossRef\]](#)
44. Kain, J.S. The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteorol.* **2004**, *43*, 170–181. [\[CrossRef\]](#)
45. Liu, S.J.; Wang, M.; Yi, X.; Shao, S.B.; Zheng, Y.Q.; Zeng, X.M. Assessing the Impact of Cumulus Parameterization Schemes on Simulated Summer Wind Speed over Mainland China. *Atmosphere* **2022**, *13*, 617. [\[CrossRef\]](#)
46. Mi, L.; Shen, L.; Han, Y.; Cai, C.S.; Zhou, P.; Li, K. Wind Field Simulation Using WRF Model in Complex Terrain: A Sensitivity Study with Orthogonal Design. *Energy* **2023**, *285*, 129411. [\[CrossRef\]](#)
47. Hong, S.-Y.; Lim, J.-O.J. The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.* **2006**, *42*, 129–151.
48. Nakanishi, M.; Niino, H. Development of an improved turbulence closure model for the atmospheric boundary layer. *J. Meteor. Soc. Jpn.* **2009**, *87*, 895–912. [\[CrossRef\]](#)
49. Siuta, D.; West, G.; Stull, R. WRF Hub-Height Wind Forecast Sensitivity to PBL Scheme, Grid Length, and Initial Condition Choice in Complex Terrain. *Weather Forecast.* **2017**, *32*, 493–509. [\[CrossRef\]](#)
50. Yu, E.; Bai, R.; Chen, X.; Shao, L. Impact of Physical Parameterizations on Wind Simulation with WRF V3.9.1.1 under Stable Conditions at Planetary Boundary Layer Gray-Zone Resolution: A Case Study over the Coastal Regions of North China. *Geosci. Model Dev.* **2022**, *15*, 8111–8134. [\[CrossRef\]](#)
51. Lee, J.C.Y.; Lundquist, J.K. Evaluation of the Wind Farm Parameterization in the Weather Research and Forecasting Model (Version 3.8.1) with Meteorological and Turbine Power Data. *Geosci. Model Dev.* **2017**, *10*, 4229–4244. [\[CrossRef\]](#)
52. Mlawer, E.J.; Taubman, S.J.; Brown, P.D.; Iacono, M.J.; Clough, S.A. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.* **1997**, *102*, 16663–16682. [\[CrossRef\]](#)
53. Dudhia, J. Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.* **1989**, *46*, 3077–3107. [\[CrossRef\]](#)
54. Benjamin, S.G.; Grell, G.A.; Brown, J.M.; Smirnova, T.G. Mesoscale weather prediction with the RUC hybrid isentropic-terrain-following coordinate model. *Mon. Weather Rev.* **2004**, *132*, 473–494. [\[CrossRef\]](#)
55. Kamath, A.; Poojari, S.; Varsha, K. Assessing the Robustness of Normality Tests under Varying Skewness and Kurtosis: A Practical Checklist for Public Health Researchers. *BMC Med. Res. Methodol.* **2025**, *25*, 206. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Liao, W.; Fang, J.; Ye, L.; Bak-Jensen, B.; Yang, Z.; Porte-Agel, F. Can we trust explainable artificial intelligence in wind power forecasting? *Appl. Energy* **2024**, *376*, 124273. [\[CrossRef\]](#)
57. Au, Q.; Herbringer, J.; Stachl, C.; Bischl, B.; Casalicchio, G. Grouped feature importance and combined features effect plot. *Data Min. Knowl. Discov.* **2022**, *36*, 1401–1450. [\[CrossRef\]](#)
58. Dong, W.; Sun, H.; Tan, J.; Li, Z.; Zhang, J.; Zhao, Y.Y. Short-term regional wind power forecasting for small datasets with input data correction, hybrid neural network, and error analysis. *Energy Rep.* **2021**, *7*, 7675–7692. [\[CrossRef\]](#)
59. Lin, W.-H.; Wang, P.; Chao, K.-M.; Lin, H.-C.; Yang, Z.-Y.; Lai, Y.-H. Wind power forecasting with deep learning networks: Time-series forecasting. *Appl. Sci.* **2021**, *11*, 10335. [\[CrossRef\]](#)
60. Shah, A.A.; Aftab, A.A.; Han, X.; Baloch, M.H.; Honnurvali, M.S.; Chauhdary, S.T. Prediction error-based power forecasting of wind energy system using hybrid WT-ROPSO-NARMAX model. *Energies* **2023**, *16*, 3295. [\[CrossRef\]](#)
61. Karaman, Ö.A. Prediction of wind power with machine learning models. *Appl. Sci.* **2023**, *13*, 11455. [\[CrossRef\]](#)
62. Tang, Y.; Yang, K.; Zheng, Y.; Ma, L.; Zhang, S.; Zhang, Z. Wind power forecasting: A transfer learning approach incorporating temporal convolution and adversarial training. *Renew. Energy* **2024**, *224*, 120200. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.