ISSN: 2688-8645

Short-Term Wind Power Forecasting Method Based on Implicit Scenario Discovery and Dynamic Weighted Fusion

Jingwei Ren [1], Fengru Ding [2], Lili Wei [2]

[1] College of Law and Political Science, North China Electric Power University, Baoding, Hebei, 071003, P.R. China

[2] College of Control and Computer Engineering, North China Electric Power University, Baoding, Hebei, 071003, P.R. China

Abstract

Wind power intermittency and uncertainty significantly increase the complexity of power grid dispatch and impose more stringent requirements on the secure and stable operation of power systems. To address this challenge, this paper proposes a scenario-oriented dynamic selection and fusion decision framework based on dynamic quantile thresholds and a performance matrix. First, a forecasting model pool consisting of physical mechanism models, statistical regression models, and individual deep learning models is constructed, and the applicable boundaries of different model categories are clarified. Second, a three-dimensional orthogonal feature space is established, and dynamic quantile thresholds are introduced to enable implicit scenario identification for wind power time series. Finally, a pyramidal evaluation index system is developed, together with a dynamic selection and fusion triggering mechanism. On the basis of the scenario–model performance matrix, the proposed framework adaptively switches between the two modes of single-model locking and multi-model weighted fusion. Experimental results using measured data from a wind farm show that the proposed method effectively mitigates the degradation in forecasting performance of individual models under extreme scenarios, such as ramping events and cut-in/cut-out moments. Consequently, the method significantly improves forecasting accuracy while demonstrating strong scenario adaptability and robustness.

Keywords

Implicit Scenarios; Power Forecasting; Weighted Fusion; Dynamic Selection.

1. Introduction

Wind power, as an important form of clean and low-carbon energy, is playing an increasingly vital role in the energy transition and the development of new-type power systems. Wind farms convert captured wind energy into electrical energy and feed it into the grid to achieve wind power generation[1]. However, wind power output is influenced by multiple factors, such as wind speed, wind direction, ambient temperature, atmospheric pressure, and complex terrain, resulting in pronounced randomness, volatility, and non-stationarity. These characteristics pose longstanding challenges to wind power forecasting, including difficulties in feature extraction, inadequate model generalization, and limited forecasting accuracy. Accurate wind power forecasting is not only essential for ensuring the secure and stable operation of power grids and improving renewable energy accommodation, but also provides critical support for the optimal dispatch and economic operation of wind farms.

Traditional wind power forecasting approaches mainly include physical modeling methods, statistical analysis methods, and machine-learning-based methods. Physical methods generally rely on meteorological information and turbine operation mechanisms and are therefore highly

ISSN: 2688-8645

interpretable. However, under complex terrain and highly variable meteorological conditions, they often face difficulties in simultaneously achieving high accuracy and broad applicability. Statistical methods, although relatively simple to construct, rely strongly on the quality of historical data and linear assumptions, which limits their ability to characterize the complex nonlinear evolution of wind power series. With the rapid development of artificial intelligence, deep learning methods have gradually emerged as an important direction in wind power forecasting research because of their strengths in nonlinear mapping, time-series modeling, and feature representation. In particular, to address the two core issues of optimization difficulty caused by the dependence of model parameters on manual experience and inaccurate feature extraction caused by the non-stationary nature of wind power data, related studies have increasingly focused on intelligent optimization, sequence decomposition, and multi-model fusion.

In terms of intelligent parameter optimization and multi-unit forecasting, Reference [2] have proposed a hybrid wind-turbine cluster power forecasting model integrating DBSCAN clustering and the enhanced hunter-prey optimization algorithm (ENHPO). By using the Pearson correlation coefficient to screen key variables affecting wind power, and employing ENHPO to optimize the DBSCAN clustering parameters so as to improve turbine grouping performance, these studies further select highly correlated representative wind turbines from each group and combine them with a multivariate long short-term memory network for power forecasting, thereby balancing both the accuracy and efficiency of large-scale wind-turbine cluster power prediction. Reference [3] focused on the strong randomness of wind power and the limited accuracy of traditional models, and proposed a method that combines the Grey Wolf Optimizer with Variational Mode Decomposition. Furthermore, the Rime Optimization Algorithm was employed to optimize the learning parameters of the Long Short-Term Memory network, thereby enhancing the model's ability to capture the features of wind power sequences. The paper [4] proposes a hybrid model for short-term wind power forecasting by integrating an improved dung beetle optimizer, variational mode decomposition, a temporal convolutional network, a gated recurrent unit, and an attention mechanism, with experimental results demonstrating significantly higher forecasting accuracy than multiple benchmark models. Furthermore , Reference [5] integrated intelligent optimization, data decomposition, and long short-term memory networks into a unified forecasting framework, and employed the butterfly optimization algorithm to jointly optimize the decomposition parameters and network parameters, enabling the model to achieve strong forecasting accuracy and generalizability under different seasonal scenarios. Reference [6] further incorporated an attention-weighted environmental factor mechanism to construct a two-stage wind power forecasting framework, in which the original sequence and the error sequence were modeled separately. By integrating environmental information into the error correction process, the proposed model improved its ability to capture both the trend and fluctuation characteristics of wind power output.

Apart from the parameter optimization issue, the strong nonlinearity, non-stationarity, and high-noise characteristics of wind power series are also important factors that constrain further improvements in forecasting performance. In response to this problem Reference [7] constructed a wind power forecasting model integrating Singular Spectrum Analysis, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, Fuzzy Entropy Reconstruction, an improved long-sequence forecasting model, and a Random Vector Functional Link network without direct links. Through sequence decomposition, feature reconstruction, and component-wise forecasting, the model improved its ability to characterize complex fluctuation patterns and enhanced forecasting accuracy. Reference [8] proposed a wind power prediction method that combines smoothing filtering with a parallel deep architecture, in which the input wind speed series was denoised and smoothed using the Savitzky-Golay filter, while the

ISSN: 2688-8645

Temporal Convolutional Network and the Long Short-Term Memory neural network were integrated in parallel to improve forecasting performance while balancing model complexity and computational efficiency. Additionally, Reference [9] proposed a process time-series Transformer model for wind power prediction, in which noise filtering was performed using the exponentially weighted moving average method, seasonal trend decomposition and a temporal autoencoder were employed to extract complex spatiotemporal features, and a sparse attention mechanism was introduced to enhance the representation capability of high-dimensional latent features, thereby improving the accuracy and stability of short-term multi-step wind power forecasting. Similarly reference [10] approached short-term wind power prediction from the perspective of data quality control and proposed a soft sensor model that integrates data preprocessing, variational mode decomposition, and long short-term memory (LSTM), in which isolation forest was used for anomaly detection, while missing value imputation and sequence decomposition-based denoising were employed to improve model convergence and forecasting accuracy.

In the field of deep learning model architecture design, relevant studies are no longer limited to the local optimization of a single model, but have gradually shifted toward an integrated development path characterized by collaborative modeling with sequence decomposition, attention mechanisms, spatiotemporal feature extraction, and hybrid architectures Reference [11] proposed a combined model for wind power forecasting by integrating the Convolutional Block Attention Module and Long Short-Term Memory, in which the attention module was used to extract wind power time-series features and the spatial features contained in numerical weather prediction data, thereby improving forecasting performance Compared with studies emphasizing multi-model combination strategies, Reference [12] proposed an ultra-short-term wind power prediction combination model based on a multilayer perceptron, a bidirectional long short-term memory network, and a temporal convolutional network, with the Pearson correlation coefficient and linear programming used to improve forecasting accuracy and reliability Notably, beyond modeling based solely on historical power series, some studies have attempted to extend the applicability of wind power forecasting methods from the perspective of large-scale meteorological information and spatial downscaling. Reference [13] proposed a computational fluid dynamics-based wind field downscaling method to overcome the limited spatial resolution of numerical weather prediction, enabling local high-resolution wind field forecasting for wind farms.

However, existing methods fail to address the limitations of traditional scenario partitioning based on fixed thresholds, which are inadequate for capturing the seasonal non-stationary variations of wind resources. Furthermore, the “global average” evaluation scheme tends to mask differences in model performance under diverse wind conditions. To overcome these limitations, this paper proposes an implicit scenario discovery and multidimensional evaluation framework based on dynamic quantile thresholds. By introducing a three-dimensional orthogonal feature space and an adaptive thresholding algorithm, the proposed framework effectively remedies these shortcomings.

This paper is devoted to improving power forecasting accuracy under complex wind conditions and overcoming the limitations of the traditional single-model prediction paradigm. To this end, a scenario-oriented dynamic selection and fusion decision framework based on dynamic quantile thresholds and a scenario-model performance matrix is proposed. First, within a Bayesian inference framework and under data-driven principles, physical mechanism models, statistical regression models, and individual deep learning models are developed. Their applicability boundaries are identified for the low-wind-speed steady-state regime and the high-wind-speed highly fluctuating regime. Second, a three-dimensional orthogonal feature space is constructed. Dynamic quantile thresholds are introduced to enable implicit scenario discovery. This maps non-stationary wind power series into identifiable discrete scenario

ISSN: 2688-8645

labels. On this basis, a pyramid-shaped evaluation system is established, incorporating basic accuracy, extreme risk, and trend-capturing capability. Finally, a dynamic selection and fusion triggering mechanism is proposed. Guided by the scenario-model performance matrix, this mechanism adaptively switches between single-model locking and multi-model weighted fusion. The fusion process is implemented through a three-dimensional dynamic weighting algorithm integrating basic accuracy, diversity penalty, and time-segmented local adjustment, together with Huber loss correction and physical boundary constraints, thereby effectively mitigating the prediction collapse of single models under specific extreme scenarios, such as ramp events and cut-in/cut-out moments.

2. Fundamental Principles of the Models

2.1. Physical-Mechanism and Statistical-Regression Models

Physical mechanism and statistical regression models are intended to capture the explicit physical laws underlying the wind power conversion process. Within the Bayesian inference framework, such models assume that the observed data follow a specific physically informed prior distribution. Under "steady-state wind conditions," characterized by smooth wind speed variations and low turbulence intensity, turbine power is primarily governed by aerodynamic equations. In this case, the bias of physical models mainly arises from the systematic errors of numerical weather prediction (NWP), rather than from the model structure itself. These algorithms are typically developed on the basis of Blade Element Momentum (BEM) theory or simplified power curve mapping.

These algorithms are typically developed on the basis of Blade Element Momentum (BEM) theory or simplified power curve mapping.

Input layer: Numerical Weather Prediction data: $X_{nwp}=\{v_{nwp},\theta_{nwp},T_{nwp}\}$, including variables such as wind speed, wind direction, and temperature.

2. Physical Mapping Layer: A coordinate transformation is carried out based on the theoretical wind turbine power curve P_{curve}. Assuming ideal operating conditions, the theoretical output power P_{theo} represented by a piecewise function with respect to wind speed v

$$P_{theo}(v)=\begin{cases}0, v\le v_{in}\\ P_{rated}\cdot\left(\frac{v-v_{in}}{v_r-v_{in}}\right)^{\alpha}, v_{in}<v\le v_r\\ P_{rated}, v_r<v\le v_{out}\\ 0, v>v_{out}\end{cases} \quad (1)$$

In Equation (1): v_{in},v_r,v_{out} represent the cut-in, rated, and cut-out wind speeds, respectively, while α denotes an empirical exponent and P_{rated} is the rated power.

3. Correction and output: To address the discrepancy between the theoretical curve and actual operation, a statistical regression correction term. ϵ_{reg} (such as linear regression or polynomial regression) is introduced, to obtain the final predicted value $\hat{y}_A$:

$$\hat{y}_A=f_{physics}(X_{nwp})+\epsilon_{reg} \quad (2)$$

where $f_{physics}$ denotes the physical mapping function based on the theoretical power curve. Reference [14] have shown that incorporating physical constraints into deep learning frameworks can significantly improve forecasting reliability. For instance, Mclean et al. proposed bounded Gaussian process models to ensure that power curve predictions remain within physically possible ranges, thereby increasing operator confidence in model outputs.

ISSN: 2688-8645

The algorithm demonstrates excellent performance in the low-wind-speed stable regime and possesses strong physical interpretability. However, when wind conditions enter a highly volatile regime or ramp events occur, the model often exhibits significant lag and underfitting because it neglects the nonlinear dissipation processes of atmospheric turbulence as well as the dynamic coupling of wake effects.

2.2. Single Deep Learning and Ensemble Learning Model

Individual deep learning models and ensemble learning models are developed to address the complex fitting problem associated with nonlinear mappings. In contrast to the “white-box” physical assumptions underlying physical mechanism models and statistical regression models, these approaches are essentially “black-box” data-driven models. Their fundamental assumption is that historical data encode statistical dependencies relevant to future states. By means of high-dimensional nonlinear transformations, these models are able to capture implicit relationships between input features and power output. In this study, two representative models are selected as candidates for the individual deep learning and ensemble learning models:

XGBoost [15] : Based on the Additive Model and the forward stage-wise algorithm, this method iteratively fits the residual distribution. It is particularly effective in handling discrete features and nonlinear decision boundaries, and can effectively capture complex feature interactions in the medium-wind-speed nonlinear regime. he objective function is defined as

$$\mathcal{L} = \sum_i l\,(y_i, \hat{y}_i) + \sum_k \Omega\,(f_k) \tag{3}$$

In Equation (3), l denotes the loss function, and Ω denotes the regularization term.

Long Short-Term Memory (LSTM) [16]:By introducing gated units (input, forget, and output gates), LSTM addresses the vanishing gradient problem in traditional RNNs.It is capable of capturing long-term temporal dependencies, and responds more rapidly to sudden wind speed changes than static models in the high-wind-speed volatile regime.The state update process is shown as follows.

$$h_t = LSTM(x_t, h_{t-1}) \tag{4}$$

In Equation (4) h_t denotes the hidden state vector, x_t is the input feature vector at time t, and h_{t-1} is the hidden state vector at the previous time step.

Although individual deep learning and ensemble learning models may achieve higher accuracy than physical mechanism and statistical regression models in specific scenarios, they remain constrained by a scenario-dependency dilemma. A standalone XGBoost or LSTM model is susceptible to a “local optimum trap”: XGBoost responds sluggishly to abrupt temporal transition points, whereas LSTM tends to overfit and generate spurious oscillations under prolonged steady-state conditions. If a single intelligent model is imposed across all scenarios, both prediction variance and prediction bias are likely to increase simultaneously. Recent studies have shown that ensemble or fusion strategies can effectively mitigate this issue [17].

3. Scenario-Based Dynamic Model Selection and Fusion Decision Framework

The scenario-based dynamic model selection and fusion decision framework is not a single mathematical model, but an intelligent system organized around four key stages: "perception-evaluation-decision-correction". This architecture aligns with recent advances in physics-informed spatio-temporal networks, where dynamic feature adaptation and physical penalty losses are employed to enhance prediction accuracy [18]. Specifically, the proposed algorithm performs implicit scenario discovery based on dynamic quantile thresholds, constructs a scenario-model performance matrix, and dynamically determines whether the “optimal single

ISSN: 2688-8645

model" or "multi-model fusion" should be adopted according to scenario characteristics. Finally, a robust correction mechanism is incorporated to ensure physical feasibility. The specific process is illustrated in Figure 1.

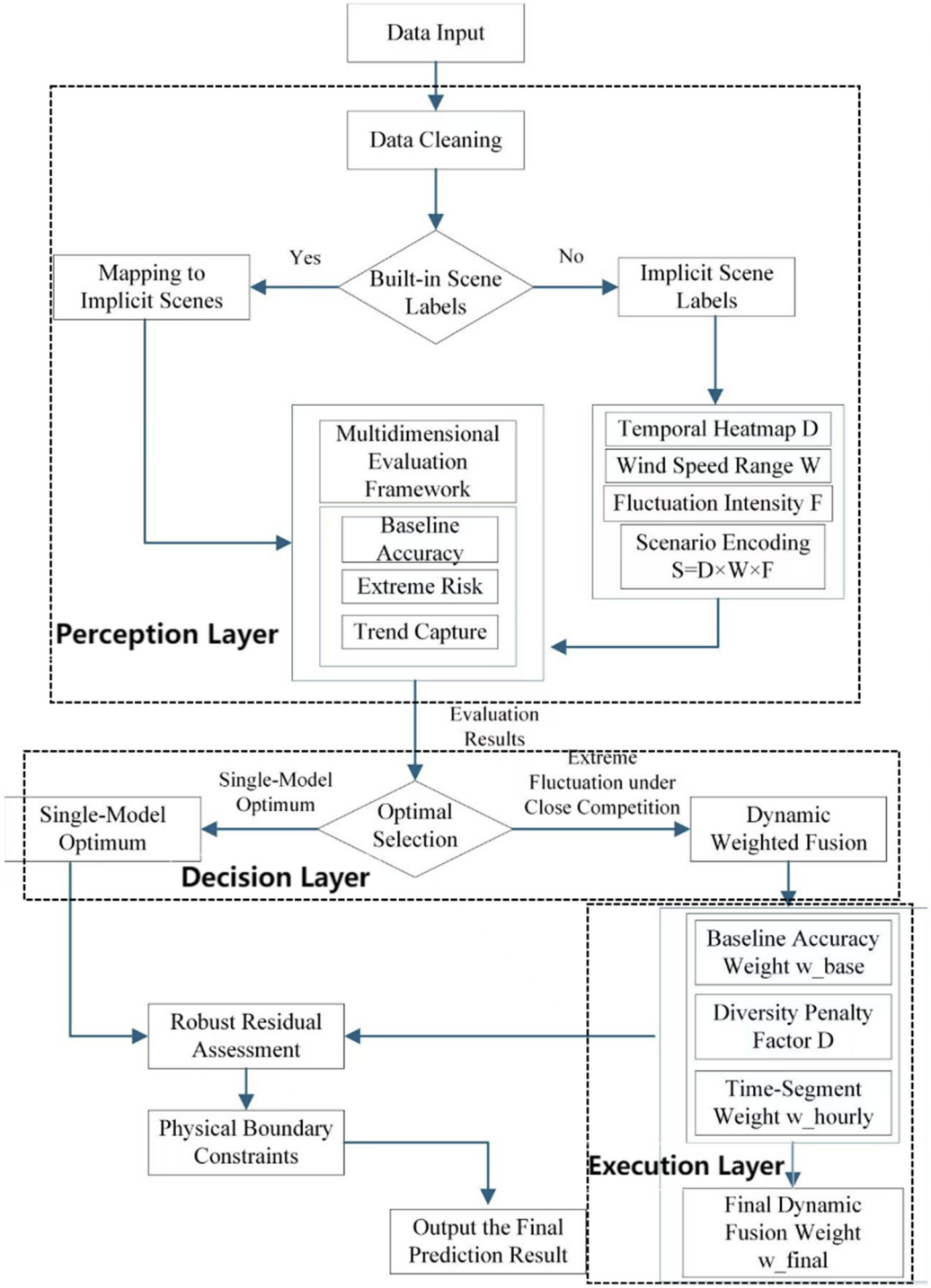

Figure 1. Flowchart of scenario-based dynamic selection and fusion decision framework

3.1. Data Input and Cleaning

As illustrated in Figure 1, the framework initiates with the Data Input and Data Cleaning stages to ensure data integrity and quality.

ISSN: 2688-8645

Data Input: The system ingests multi-source heterogeneous data, including Numerical Weather Prediction (NWP) data, real-time wind speed measurements, and historical power output series.

Data Cleaning: Raw data often contain noise, missing values, or physically impossible entries (e.g., power output when wind speed exceeds cut-out thresholds). **S**pecifically, the cleaning process focuses on removing outliers caused by sensor failures and eliminating non-zero power readings during cut-in/cut-out moments, which are physically invalid.A preliminary cleaning process is applied to remove outliers and impute missing values, thereby constructing a robust dataset for subsequent scenario discovery.

3.2. Scenario Discovery and Multidimensional Evaluation Framework (Perception Layer)

In the scenario-based dynamic model selection and fusion decision framework, the system does not directly perform forecasting; rather, it first evaluates the current wind conditions. Depending on whether the input data contain predefined scenario labels, a dual-channel processing mechanism is adopted. For labeled explicit scenarios, such as historically manually annotated "typhoon passage" events or "severe convective warnings" issued by the observatory, the corresponding labels are directly used as scenario encodings and passed to the subsequent scenario–model performance matrix retrieval and model selection procedure, thereby fully leveraging prior knowledge. For unlabeled implicit scenarios, the system instead proceeds to the implicit scenario discovery process described below, in which scenario labels are automatically generated using dynamic quantile thresholds and a three-dimensional orthogonal feature space. Ultimately, both channels output unified standardized scenario encodings, which are transparent to the downstream evaluation and decision modules. A three-dimensional orthogonal feature space is constructed, and the scenario space is defined as

$$\mathbf{S} = \mathbf{D} \times \mathbf{W} \times \mathbf{F} \tag{5}$$

By constructing a three-dimensional orthogonal feature space, a unique scenario encoding t is assigned to each time point（such as：day_high_intence）, All subsequent indicators are aggregated and calculated according to this scenario label. The use of dynamic quantile thresholds for scenario discovery is motivated by recent work on scenario generation that considers both statistical and temporal characteristics of wind power series [19].

In Equation (5)：D（temporal thermal dimension）：day/night (based on atmospheric thermal stability), where 6≤hour(t)≤18 to daytime, and otherwise to nighttime.

W is wind speed operating interval：low/medium/high (based on the physical characteristics of the wind turbine power curve){ low：$v_t < 5m/s$ mediu：$5m/s \le v_t < 10m/s$ high：$v_t \ge 10m/s$ }

F（dynamic fluctuation intensity）：Fixed thresholds are discarded in favor of dynamic quantile thresholds, which are used to adaptively determine the fluctuation thresholds, thereby classifying fluctuation intensity into three levels: stable, moderate, and intense.

A multi-level evaluation framework is established as a three-layer pyramid composed of basic accuracy, extreme risk, and trend-capturing capability, and is subsequently employed to assess the performance of each model under each scenario.

Basic accuracy layer: it measures the direct deviation between predicted values and actual values and serves as the benchmark for evaluation.

RMSE（Root Mean Square Error）：

$$RMSE = \sqrt{\frac{1}{M}\sum(P_{pred} - P_{true})^2} \tag{6}$$

ISSN: 2688-8645

MAE (Mean Absolute Error) :

$$MAE = \frac{1}{M}\sum_{i=1}^{M} | P_{pred}^{i} - P_{true}^{i} \quad (7)$$

Accuracy:

$$Accuracy = 1 - \frac{MAE}{C_{cap}} \quad (8)$$

In Equation (6) : M denotes the total number of samples, P_{pred} represents the predicted power value, P_{true} represents the actual power value.Since this metric is sensitive to large deviations, it is employed to identify whether severe extreme prediction errors are present in the forecasting results. Smaller values indicate better performance.

In Equation (7) : M denotes the total number of samples, P_{pred}^{i} represents the i-th predicted value, P_{true}^{i} represents the i-th true value.

In Equation (8) : C_{cap} denotes the installed capacity. This normalized metric expresses the error in percentage terms, thereby conforming to grid evaluation standards. Larger values indicate better performance.

The extreme risk layer emphasizes model performance under extreme weather conditions, such as strong winds and ramp events. Model performance is evaluated from two perspectives: the dynamic fluctuation capture rate and the false alarm rate at cut-in/cut-out wind speeds.

The dynamic fluctuation capture rate is defined based on the fluctuation-intensity dimension of the three-dimensional orthogonal feature space. It measures whether the predicted values exhibit synchronized drastic changes under scenarios classified as intense fluctuations, thereby assessing whether the model can promptly track abrupt wind speed variations or exhibits a delayed response.

False alarm rate at cut-in/cut-out wind speeds : this metric focuses on the prediction errors at moments ,when $v_t > 3m/s$ (cut-in) and $v_t > 20m/s$ (cut-out), and is used to determine whether the model produces spurious power output predictions under zero-power or power-limited operating conditions.

The trend-capturing layer assesses the model's ability to identify the direction of power variations from two perspectives, namely directional accuracy and ramp event capture rate, rather than focusing solely on the magnitude of the predicted values.

Directional Accuracy:

$$DA = \frac{1}{M}\sum \delta\left(sign(\Delta P_{pred}) == sign(\Delta P_{true})\right) \quad (9)$$

In Equation (9), DA represents the directional accuracy, which quantifies the model's ability to correctly predict the trend of wind power changes. δ is an indicator function that equals 1 when the condition inside the parentheses is satisfied and 0 otherwise,$sign$ is the sign function, which returns 1 for positive values, -1 for negative values, and 0 for zero, to judge the direction of power variation. ΔP_{pred} and ΔP_{true} are the predicted and actual power differences between two adjacent time steps, respectively, which are used to verify the consistency between the predicted and real power trends.

3.3. Dynamic Model Selection and Fusion Triggering Mechanism (Decision Layer)

Based on the evaluation framework presented in Section 2.4.1, an N×K performance matrix denoted by M_{eff} is constructed, where N represents the number of scenarios and K represents the number of candidate models.According to the real-time scenario label L_t,the system queries the corresponding position in the matrix and performs the following operational procedure.

ISSN: 2688-8645

(1) Adaptive single-model selection: if a significantly superior model exists in the current scenario, with a score substantially higher than those of the other models, the system directly selects and locks onto this model:

$$M_{current} = arg\ max\left(M_{eff}[L_t, j]\right) \tag{10}$$

In Equation (10), $M_{current}$ is obtained by selecting the index j of the model with the maximum performance score in the current scenario row L_t of matrix M_{eff}.

(2) Multi-model fusion: Multi-model fusion is triggered under two modes. Performance-tie triggering: if the score difference between the best model and the second-best model is smaller than a threshold $\Delta S < \theta_{switch}$ (0.05), indicating that no model has an absolute advantage, thus fusion is triggered to prevent misjudgment. Extreme-fluctuation scenario triggering: if the real-time wind condition is detected to belong to the "intense fluctuation" category (with the standard deviation exceeding the P_{67} percentile), **fusion is forcibly activated to improve robustness.**

3.4. Three-dimensional Dynamic Weighted Fusion Algorithm (Execution Layer)

When the multi-model fusion mechanism described in Section 2.2 is triggered, the system not simply adopts a simple arithmetic average. Instead, it calculates three-dimensional dynamic weights.

Step 1: Basic accuracy weight. Based on the inverse-square allocation of historical RMSE, the system strengthens the preference for high-accuracy models and penalizes low-accuracy models.

$$w_{base,i} = \frac{1}{RMSE_i^2 + \epsilon} \tag{11}$$

In Equation (11): $w_{base,i}$ denotes the basic weight of model i .$RMSE_i$ represents the root mean square error of model i on the historical validation set.ϵ is a very small constant, introduced to prevent division by zero.

Step 2: Diversity penalty factor. To prevent redundancy, that is, multiple models making similar errors, a penalty mechanism based on the Spearman rank correlation coefficient is introduced. Calculate the average absolute correlation $\bar{\rho}_i$, between model i and all the other models.

Apply a piecewise penalty function:

$$D_i = \begin{cases} 0.5 & if\ \bar{\rho}_i > \tau \\ 1.0 - 0.5 \cdot \bar{\rho}_i & otherwise \end{cases} \tag{12}$$

In Equation (12): D_i denotes the diversity penalty factor of model i ,and $\bar{\rho}_i$ represents the mean correlation between the prediction sequence of model i and those of the other models.τ is the correlation threshold (e.g. 0.8). If the models are highly similar, the weight is forcibly reduced by 50%; if a model is relatively distinctive, with strong complementarity, most of its weight is retained.

Step 3: Time-segment-based local weight adjustment.This step takes intraday periodicity into account and performs local optimization for the current forecasting time h

$$w_{hourly,i} = \frac{1}{Mean(RMSE_{i,h}) + \epsilon} \tag{13}$$

In Equation (13), $w_{hourly,i}$ denotes the local weight of model i at the current hour h. $Mean(RMSE_{i,h})$ represents the average RMSE of model i over the same hourly period in historical data.

Step 4: Comprehensive weight synthesis. The weights from the above three dimensions are normalized to obtain the final fusion weight vector W_{final}.

ISSN: 2688-8645

$$w'_i = w_{base,i} \times D_i \times w_{hourly,i} \quad (14)$$

$$w_{final,i} = \frac{w'_i}{\sum_{j=1}^{K} w'_j} \quad (15)$$

In Equation (15): $w_{final,i}$ denotes the final dynamic fusion weight of model i reflecting its overall reliability under the current time and scenario.

3.5. Output Correction

To prevent extreme deviations caused by sudden gusts or sensor failures, the preliminary fused value,$\widehat{P}_{\text{fusion}}$ is further corrected

The value of $\widehat{P}_{\text{fusion}}$can be calculated using $w_{final,i}$ as follows:

$$\hat{P}_{fusion} = \sum_{i=1}^{K} w_{final,i} \times P_i \quad (16)$$

In Equation (16): p_i denotes the original predicted value of the i-th model and $w_{final,i}$ represents the weight of the i-th model in the final result,K denotes the total number of candidate models included in the fusion. $\hat{P}_{fusion}$ is the fused predicted value, which serves as the input to the robust correction stage.

Residual analysis: calculate the standard deviation of historical residuals σ.

Dynamic threshold determination: set the threshold as T =$\delta \cdot \sigma$(where δ is usually taken as 1.35).

Piecewise correction strategy: if $|r| \leq T$:the residual is regarded as normal noise and no correction is applied. If $|r| > T$:it is regarded as an abnormal spike, and linear compression correction is performed as follows:

$$\hat{P}_{corrected} = \hat{P}_{fusion} - sign(r) \cdot (|r| - T) \cdot \alpha \quad (17)$$

In Equation (17) : α denotes the correction coefficient, which controls the compression strength.

3.6. Physical Boundary Constraints

The final step of this algorithm is intended to ensure that the mathematical results are consistent with engineering reality. The importance of physical constraints in wind power forecasting has been emphasized in recent research, where physical penalty terms are incorporated into loss functions to guarantee that prediction outputs adhere to physical laws [20].

$$\hat{P}_{final} = Clip(\hat{P}_{corrected}, 0, P_{rated}) \quad (18)$$

In Equation (18) : P_{rated} denotes the rated power (installed capacity) of the wind turbine. The Clip function forcibly constrains the calculated value within a reasonable range, i.e., Clip (value, max, min) .At the same time, the ramp rate is also checked to ensure that the power variation does not exceed the physical limit.

4. Case Study Analysis

4.1. Experimental Data

This study utilizes actual time-series operational data from a domestic wind farm in 2022, with a temporal resolution of 15 minutes. The dataset encompasses three core dimensions: equipment identification, operational status, and meteorological conditions. As detailed in Table 1, the original data underwent engineering preprocessing, including the removal of outliers (e.g., extreme values and non-zero power readings when wind speeds exceeded the cut-out threshold). The dataset, comprising 26,499 samples, was partitioned into training, validation, and test sets in a 7:2:1 ratio. Multi-source data were precisely aligned along the time

ISSN: 2688-8645

axis for model training, parameter tuning, and performance verification to ensure the reproducibility and engineering generalization capability of the results.

Table 1. Data Content

Data Dimension	*Specific Content*
Equipment ID & Time Dimension	*Wind turbine ID, Timestamp*
Operation Data Dimension	*Three-phase measured active power, Power prediction deviation, Nacelle temperature*
Meteorological Data Dimension	*Measured wind speed, Wind direction, Ambient temperature, Air pressure, Forecast wind direction corresponding to Numerical Weather Prediction (NWP)*

4.2. Comparison of Model Prediction Results

To verify the effectiveness of each module in the proposed scheme, multiple comparative experiments are conducted, covering two categories of benchmarks: single models and fusion methods. Specifically, the single prediction models include the physical mechanism benchmark model (Physical), XGBoost model, and LSTM model. The fusion model proposed in this paper is a Physical+XGBoost+LSTM dynamic weight fusion model (Fusion), which incorporates a diversity penalty term and Huber robust correction mechanism to enhance the model's generalization and noise immunity.

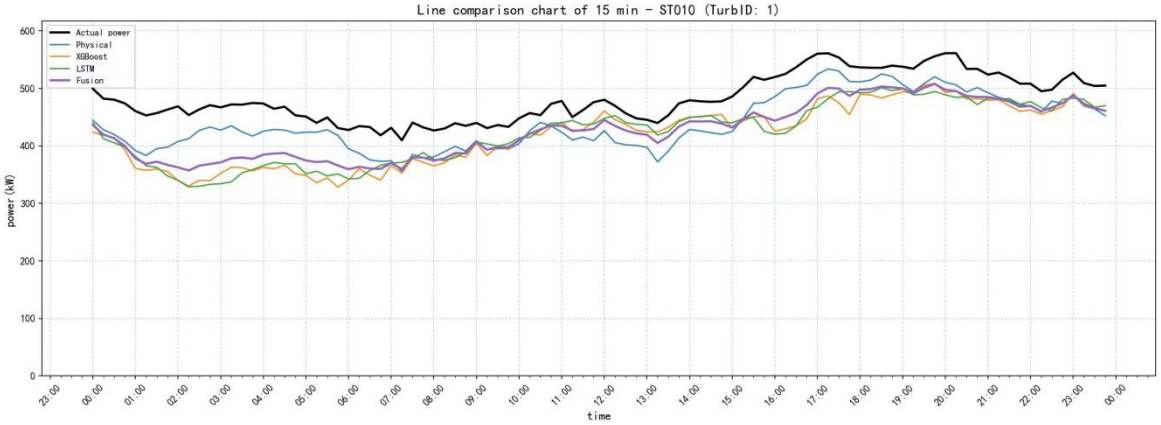

Figure 2. 24-hour wind power prediction results of wind turbine 24

Figure 2 shows the 24-hour wind power prediction results of Wind Turbine 1 at a 15-minute time resolution, clearly comparing the prediction performance of three approaches: the physical mechanism benchmark model (Physical), data-driven models (XGBoost, LSTM), and the proposed fusion model (Fusion). The physical mechanism benchmark model is constructed based on physical laws. However, affected by the uncertainty of meteorological input, its prediction results are generally conservative, with highly random curve fluctuations and poor consistency with actual power values. Particularly during periods of high-frequency wind speed fluctuations (e.g., 11:00–13:00), its response to power mutations lags behind other models, failing to effectively capture the rapid dynamic characteristics of wind power.

Among them, as a tree-based data-driven model, XGBoost exhibits excellent accuracy in local nonlinear fitting scenarios but tends to fall into local optima. During stable power periods (e.g., 04:00–06:00), it shows obvious underestimation and lag, resulting in a slow response to time-series mutations.

ISSN: 2688-8645

As a deep learning time-series data-driven model, LSTM is capable of capturing long-term temporal dependencies but is prone to overfitting under steady-state conditions, leading to physically meaningless false oscillations.

The Fusion model achieves an optimal balance between global prediction error and robustness through the complementary advantages of multiple models.This comes at the cost of moderate compromise in local short-term accuracy. Compared with single models, its prediction curve demonstrates superior fitting degree and trend consistency with the actual power. It not only effectively eliminates the response lag of the physical mechanism model but also mitigates the overfitting and false oscillations of data-driven models. Consequently, it exhibits more stable and reliable prediction performance across all operating conditions, significantly enhancing the accuracy of wind power prediction while ensuring strong stability and engineering practicality.

4.3. Experimental Verification Scheme

Two groups of comparative experiments are designed to hierarchically verify the effectiveness of the integrated closed-loop scheme of "prediction-evaluation-attribution-report" proposed in this paper. All experiments are independent and progressive.

Combined with the engineering assessment requirements of wind power prediction, the evaluation indicators of prediction performance are designed as shown in Table 2. All indicators support scenario-based aggregation calculation.

Table 2. Evaluation Indicators

Indicator Name	*Formula*	*Function*
RMSE	$RMSE = \sqrt{\frac{1}{M}\sum(P_{pred} - P_{true})^2}$	*Sensitive to large errors, suitable for scenarios focusing on extreme errors.*
MAE	$MAE = \frac{1}{M}\sum_{i=1}^{M} \mid P_{pred}^{i} - P_{true}^{i}$	*Reflects average deviation.*
Accuracy	$Accuracy = 1 - \frac{MAE}{C_{cap}}$	Complies with power grid assessment standards, has engineering practicability.

4.3.1. Precision Verification Scheme of Multi-model Fusion Framework

The global prediction performance comparison between each model and the fusion model proposed in this paper is shown in Table 3. The Fusion model is significantly superior to the single prediction model in all core indicators.The specific advantages of the Fusion model are reflected in the following two aspects: the error is significantly reduced, the RMSE is 152.41kW, the maximum reduction is 37.55% compared with the single model, the MAE is 112.30kW, and the maximum decrease is 45.74%. The accuracy is steadily improved, and the normalized accuracy reaches 92.76%, which is 0.33, 2.85 and 6.10 percentage points higher than that of XGBoost, LSTM and physical mechanism benchmark model respectively. The Physical+XGBoost+LSTM multi-model fusion framework realizes the organic combination of interpretability, nonlinear fitting and time-series dependence expression ability. Compared

ISSN: 2688-8645

with the single model, it significantly improves the wind power prediction accuracy, and the global normalized accuracy is high.

Table 3. Overall Prediction Accuracy of Each Model

Model	*RMSE (kW)*	*MAE (kW)*	*Accuracy (%)*
Fusion	***152.41***	***112.30***	***92.76***
XGBoost	*196.18*	*177.43*	*92.43*
LSTM	*235.18*	*156.46*	*89.91*
Physical	*244.04*	*206.95*	*86.66*

4.3.2. Effectiveness Verification Scheme of Implicit Scene Discovery Mechanism

Two typical scenarios, middle wind speed (5-10m/s) and high wind speed (≥10m/s), are selected to compare the normalized accuracy of physical mechanism model, XGBoost, LSTM and Fusion models in different scenarios. The results are shown in Table 4.Meanwhile, the capability of global evaluation and scenario-based evaluation in identifying model shortcomings is compared.

Table 4. Prediction Accuracy of Wind Farm Under Different Scenarios (%)

Model	*Middle Wind Speed*	*High Wind Speed*
Physical	*98.50*	*94.69*
XGBoost	*98.81*	*97.38*
LSTM	*94.02*	*94.50*
Fusion	***98.63***	***97.80***

The comparative data in Table 4 clearly reveals the performance differences of different modeling strategies in typical wind speed scenarios. The Physical achieves 98.50% prediction accuracy in the medium wind speed range, verifying its theoretical effectiveness under the mainstream operating conditions of wind turbines. However, in the high wind speed range, affected by complex factors such as turbulence and nonlinearity, the prediction accuracy drops to 94.69%, exposing the poor adaptability of the physical model to extreme working conditions. Relying on its strong nonlinear fitting ability, XGBoost improves the accuracy to 98.81% in medium wind speed scenarios, significantly corrects the deviation of the physical model, and fully demonstrates the generalization advantage of data-driven model in complex wind conditions. The accuracy of LSTM time series model under two wind speeds is 94.02% and 94.50% respectively. The overall performance is slightly lower than that of the previous two

models, but it provides complementary support for time series feature extraction for the fusion model.The Fusion model achieves 98.63% in the medium wind speed range and 97.80% in the high wind speed range. In the medium wind speed range, its performance is comparable to that of XGBoost, while in the more challenging high wind speed range, it outperforms all single models. This effectively balances the organic combination of physical mechanism theoretical constraints and data-driven generalization ability, proving the effectiveness of the fusion strategy in improving the overall robustness and full-condition generalization ability of the model.

5. Summary

This paper proposes an implicit scenario discovery and multidimensional evaluation framework based on dynamic quantile thresholds, and further constructs a scenario-driven dynamic model selection and robust fusion decision-making system, which effectively addresses the challenges of accuracy and robustness in wind power forecasting.The main contributions of this work are summarized as follows:

1) To overcome the limitation of traditional scenario classification, which relies on fixed thresholds and cannot adapt to seasonal non-stationary changes in wind resources, a three-dimensional orthogonal feature space is constructed using time, wind speed, and dynamic fluctuation intensity, thereby enabling refined segmentation of complex wind conditions.

2) A multi-level evaluation framework and a “scenario–model performance matrix” are established. By retrieving this matrix in real time, the system can determine the most suitable model under the current wind condition, thus compensating for the inability of a single model to perform well across all scenarios.

3) A dynamic model selection and fusion triggering mechanism is designed to adaptively choose between “single best model” and “multi-model fusion” according to scenario characteristics. In addition, robust correction based on Huber loss and physical boundary constraints are introduced to ensure the physical feasibility of the forecasting results.

Overall, the proposed method not only provides a solution for wind power forecasting under complex wind conditions that simultaneously achieves accuracy, robustness, and interpretability, but also offers theoretical support and methodological reference for the scenario-oriented application and intelligent decision-making of renewable energy power forecasting models.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

[1] Li G, Yu Y, Wang Y, et al. Research progress and development trend of wind farm power prediction[J]. J. Drain. Irrig. Mach. Eng., 2024, 42: 778-784.

[2] Hou G, Wang J, Fan Y. Wind power forecasting method of large-scale wind turbine clusters based on DBSCAN clustering and an enhanced hunter-prey optimization algorithm[J]. Energy Conversion and Management, 2024, 307: 118341.

[3] Wang Xiangwei. Ultra-short-term wind power prediction based on GWO-VMD-RIME-LSTM[J]. Chinese Journal of Construction Machinery, 2026, 24(1): 50-55. DOI:10.15999/ j.cnki. 3119 26. 2026.01.014.

[4] He Y,Wang W,LiM,et al.A short-term wind power prediction approach based on an improved dung beetle optimizer algorithm, variational modal decomposition, and deep learning[J]. Computers and Electrical Engineering, 2024, 116: 109182.

[5] Wang Y, Zhao K, Hao Y, et al. Short-term wind power prediction using a novel model based on butterfly optimization algorithm-variational mode decomposition-long short-term memory[J]. Applied Energy, 2024, 366: 123313.

[6] Xiao Y, Wei H, Shi Y, et al. A short-term wind power prediction based on MCOOT optimized deep learning networks and attention-weighted environmental factors for error correction[J]. Energy, 2025, 324: 136054.

[7] Jia Q, Kang K, Wang B, et al. Research on wind power prediction with secondary decomposition and multi-algorithm fusion for complex nonlinear time series[J]. Computers and Electrical Engineering, 2025, 128: 110688.

[8] Liu S, XU T, Du X, et al. A hybrid deep learning model based on parallel architecture TCN-LSTM with Savitzky-Golay filter for wind power prediction[J]. Energy Reports, 2022, 8: 1189-1202.

[9] Zhang Z, Sun Z, Guo X, et al. Short-term multi-step wind power prediction model based on Pt-Transformer neural network integrating spatio-temporal feature and sparse attention[J]. Electric Power Systems Research, 2025, 248: 111970.

[10] Lei P, Ma F, Zhu C, et al. LSTM short-term wind power prediction method based on data preprocessing and variational modal decomposition for soft sensors[J]. Sensors, 2024, 24(8): 2521.

[11] Zhang Zhe, Wang Bo. Short-term power prediction method of wind power cluster based on CBAM-LSTM[J]. Journal of Northeast Electric Power University, 2024, 44(1): 1-8. DOI:10.19718/ j.issn. 1005-2992.2024-01-0001-08.

[12] Zhang Xiaoyan, Xiang Mian, Zhu Li, et al. Ultra short term wind power prediction based on MLP-BiLSTM-TCN combination[J]. Journal of Hubei Minzu University (Natural Science Edition), 2023, 41(4): 513-519, 529. DOI:10.13501/j.cnki.42-1908/n.2023.12.013.

[13] Castorrini A, Gentile S, Geraldi E, et al. Increasing spatial resolution of wind resource prediction using NWP and RANS simulation[J]. Journal of Physics: Conference Series, 2021, 1932(1): 012016.

[14] Mclean J H, Jones M R, OConnell B J, et al, Physically meaningful uncertainty quantification in probabilistic wind turbine power curve models as a damage-sensitive feature. Structural Health Monitoring, 22(6).

[15] Chen T and Guestrin C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794).

[16] Hochreiter S and Schmidhuber J, Long short-term memory. Neural Computation, 9(8), 1735-1780.

[17] Banik R and Biswas A, Interpretable wind power forecasting with residual learning-based model. Electric Power Systems Research, 247, 111824.

[18] Aslam L, Zou R,Lin L,et al. Physics-informed spatio-temporal network with trainable adaptive feature selection for short-term wind speed prediction. Computers & Electrical Engineering, 126, 110517.

[19] You Z, Liang X, Lin L, et al, Scenario Generation for Day-ahead Wind Power Output Sequences Considering Statistical and Temporal Characteristics. Modern Electric Power. DOI: 10.19725/j. cnki. 1007-2322.2024.0191.

[20] Liu Y, Gu Y, Long Y, et al.,Research on Physically Constrained VMD-CNN-BiLSTM Wind Power Prediction. Sustainability, 2025, 17(3), 1058.