EAI Endorsed Transactions
on Energy Web
Research Article EAI.EU

A Novel Hybrid Collaborative Forecasting Method for Offshore Wind Power Based on Intelligent Optimization

Ruanming Huang[1], Chen Qian[1], Tianli Song[1], Yumeng Jiang[1], Hongyu Wen[2,*], Bing Wang[2]

[1]State Grid Shanghai Municipal Electric Power Company, Shanghai 200120, China
[2]College of Artificial Intelligence and Automation, Hohai University, Nanjing 211100, China

Abstract

INTRODUCTION: With the rapid growth of offshore wind power integration and increasing wind power penetration, accurate power forecasting has become essential for maintaining grid stability and supporting economic dispatch.
OBJECTIVES: This paper aims to develop a high-accuracy offshore wind power forecasting framework that can effectively handle noisy, non-stationary data and reduce the impact of outliers on prediction performance.
METHODS: A two-stage data cleaning procedure is first constructed by combining density-based spatial clustering of applications with noise (DBSCAN) and polynomial regression to accurately identify and correct anomalous power data. The cleaned series is then decomposed using a sequential scheme that applies complete ensemble empirical mode decomposition followed by particle-swarm-optimized variational mode decomposition, producing multiple intrinsic mode components. Each component is fed into a hybrid temporal convolutional-gated recurrent unit (TCN-GRU) network, whose hyperparameters are globally tuned using an intelligent optimization algorithm, and the component-wise forecasts are aggregated to obtain the final power prediction.
RESULTS: Simulation studies based on measured data from an offshore wind farm show that the proposed method significantly reduces forecasting errors compared with conventional forecasting models and single-stage decomposition approaches.
CONCLUSION: The results demonstrate that the proposed adaptive optimization-based composite collaborative framework effectively improves both the accuracy and robustness of offshore wind power forecasting.

Keywords: Wind Power Forecasting, Data Cleaning; Modal Decomposition, Particle Swarm Optimization, Convolutional Neural Networks.

Received on 27 September 2025, accepted on 20 December 2025, published on 27 April 2026

doi: 10.4108/ew.12728

1. Introduction

With the acceleration of the global energy transition, offshore wind power has become a strategic focus owing to its abundant resources and relatively stable output [1,2]. However, the inherent intermittency and strong randomness of wind energy lead to pronounced fluctuations in power generation, posing serious challenges to the safe and stable operation of power systems [3,4]. Consequently, the development of high-precision wind power forecasting models is critical for ensuring grid security and improving the quality of optimal dispatch decisions [5]. In practice, wind power forecasting is affected by multiple sources of uncertainty, which limit both its accuracy and robustness. On the one hand, curtailment, equipment failures, communication errors, and meteorological disturbances

* Corresponding author. Email: HyWen1018@163.com

during operation can introduce missing values and abnormal data into the SCADA measurements of wind turbines; if not properly treated, these issues directly degrade model training and prediction performance. On the other hand, wind power time series exhibit strong non-stationarity, multi-scale components, and high noise levels, as well as local coupling effects caused by turbine wake interactions and inter-turbine differences. These characteristics make it difficult to effectively extract the underlying dynamic patterns directly from the raw sequences. Therefore, systematic data cleaning prior to model development is essential. By further integrating multiscale signal decomposition with adaptive forecasting models, data quality and feature separability can be enhanced, thereby improving the reliability and accuracy of wind power forecasting [6].

Regarding wind power data cleaning, substantial progress has been made in both domestic and international studies. Reference [7] adopts a combination of the quartile method and K-means clustering to remove outliers; however, the determination of the parameter k is relatively complex and may affect the stability of the results. Reference [8] develops a mathematical model based on the quartile algorithm to identify and reconstruct abnormal data. Reference [9] applies the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm for visual anomaly detection, but its ability to capture clustered anomalies remains limited. Reference [10] proposes a hybrid approach that combines median absolute deviation with the quartile method for anomaly cleaning, which effectively identifies abnormal points and achieves better overall cleaning performance than alternative methods, albeit with a slightly longer runtime compared to single-method approaches.

Currently, wind power forecasting methods are generally classified into physical models and statistical models [11]. Physical models are constructed based on numerical weather prediction and turbine mechanism descriptions, but they are difficult to formulate and strongly dependent on high-quality input data. In contrast, statistical methods establish mappings between inputs and outputs by extracting patterns from historical data; owing to their flexibility and high predictive accuracy, they have become the mainstream approach [12,13]. However, offshore wind power series exhibit pronounced non-stationarity and nonlinear behavior, making it difficult for single machine learning models to fully capture the complex coupling between wind speed and power output. Recent studies have shown that signal decomposition can effectively enhance the forecasting performance of non-stationary sequences. Reference [14] combines Variational Mode Decomposition (VMD) with a Gated Recurrent Unit (GRU) model to reduce randomness and volatility in wind power forecasting. Reference [15] performs a secondary decomposition on the highest-frequency and most disordered intrinsic mode function (IMF1) obtained from the first decomposition, further separating noise from useful high-frequency structures. This procedure produces more stationary sub-series and significantly improves prediction accuracy. Reference [16] proposes a short-term wind power forecasting method based on Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and Temporal Convolutional Networks (TCN). However, the TCN parameters must be manually tuned according to the forecasting results, making it difficult to guarantee convergence to a globally optimal solution.

In summary, the output of offshore wind turbines is strongly correlated with wind speed and historical power data. However, most existing studies focus on modeling individual time series in isolation and overlook their inherent coupling characteristics, which constrains the accuracy of forecasting models. To address data anomalies, strong non-stationarity, and mode aliasing in offshore wind power forecasting, this paper proposes a novel composite collaborative prediction method for offshore wind power based on adaptive optimization. First, a two-stage data cleaning algorithm combining DBSCAN and ninth-order polynomial regression is developed to accurately identify and correct anomalous points on the power curve, thereby improving data quality. Second, a hybrid signal decomposition strategy is introduced: the raw wind power signal is initially decomposed using CEEMDAN, and the key component containing the dominant high-frequency disturbances (IMF1) is subsequently subjected to a secondary decomposition via VMD optimized by Particle Swarm Optimization (PSO), enabling multi-scale feature separation. On this basis, a PSO-TCN-GRU prediction model is constructed, in which the dilated convolution structure of the TCN extracts multi-scale features while the GRU captures temporal dependencies. PSO is employed to automatically optimize the model hyperparameters, achieving high-precision offshore wind power forecasting. Finally, simulation experiments based on actual measurement data from an offshore wind farm demonstrate that the proposed method outperforms single-sequence decomposition approaches and traditional models in terms of MSE, MAE, and other evaluation metrics, thereby validating the effectiveness and superiority of the proposed framework.

2. Research Background Overview

2.1. DBSCAN Algorithm Principles

DBSCAN is a density-based spatial clustering algorithm that identifies arbitrarily shaped clusters and automatically detects outliers by examining the local density distribution around data points [17]. Unlike distance-based K-means, DBSCAN does not require the number of clusters to be specified in advance. Instead, it characterizes density using two key parameters: the neighborhood radius ε and the density threshold M (MinPts). Given a dataset D, the ε-neighborhood of a point p is defined as the set of points whose distance from p does not exceed εε. If the ε-neighborhood of p contains at least M points (including p itself), p is called a core point. Any point that lies within the ε-neighborhood of a core point but does not satisfy the core-point condition is termed a border point. All remaining points, which are neither core points nor border points, are classified as noise points.

The algorithm starts by randomly selecting an unvisited point. If the point is a core point, a new cluster is created and

its ε-neighborhood is recursively expanded by adding all density-reachable points until no further expansion is possible. The procedure is then repeated on the remaining unvisited points until all points in D have been processed. The DBSCAN clustering process is illustrated in Figure 1, where the ε-neighborhoods are indicated by dashed circles and the density threshold is set to $M = 6$. As shown, the ε-neighborhood of point Bcontains seven points, exceeding the threshold M, so B is a core point. Point A has fewer than six points in its own ε-neighborhood, but it lies within the $\varepsilon\varepsilon$-neighborhood of core point B, and is therefore a border point. Point C is neither a core point nor a border point, and thus is identified as a noise point.

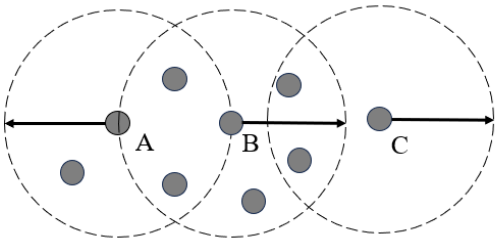

Figure 1. DBSCAN clustering process

2.2. Principles of Polynomial Regression

The power output of a wind turbine exhibits a pronounced nonlinear relationship with wind speed. A typical wind speed-power curve can be divided into four regions: start-up, ramp-up, rated, and cut-out. In the ramp-up region, the output power increases rapidly with wind speed and is commonly approximated by a polynomial function. Among various alternatives, a ninth-order polynomial is widely adopted because of its strong fitting capability. The wind speed-power relationship can therefore be expressed as:

$$P(v) = \alpha_0 + \alpha_1 v + \alpha_2 v^2 + \alpha_3 v^3 + \cdots + \alpha_9 v^9 = \sum_{k=0}^{9} \alpha_k v^k \quad (1)$$

where v denotes the wind speed, $P(v)$ is the corresponding turbine power output, $\alpha_0, \alpha_1, \cdots \alpha_9$ are the regression coefficients to be estimated.

(i) For n data points (v_i, P_i), the ninth-order polynomial regression model is given by:

$$P_i = \sum_{k=0}^{9} \alpha_k v^k + \delta_i = \hat{P}_i + \delta_i \quad (2)$$

where $\hat{P}_i$ denotes the model-predicted value corresponding to the i-th data point, and δ_i is the error term.

(ii) The model parameters are estimated using the Ordinary Least Squares (OLS) method. The objective is to find a parameter vector $\hat{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_9]^T$, that minimizes the sum of squared differences between the predicted and actual power values for all data points. The objective function is defined as:

$$S = \min \sum_{i=1}^{n} \left(P_i - \hat{P}_i\right)^2 = \min \sum_{i=1}^{n} \delta_i^2 \quad (3)$$

(iii) To evaluate the goodness of fit, the coefficient of determination on R^2 is adopted as the performance metric, which is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(P_i - \hat{P}_i\right)^2}{\sum_{i=1}^{n} \left(P_i - \bar{P}_i\right)^2} \quad (4)$$

where $\bar{P}_i$ is the arithmetic mean of all observed values P_i. A value of on R^2 closer to 1 indicates a better fitting performance.

2.3. Data Decomposition Methods

Commonly used signal decomposition methods include Variational Mode Decomposition (VMD) [18], Singular Spectrum Analysis (SSA) [19], and Empirical Mode Decomposition (EMD) [20]. VMD formulates the decomposition as a constrained variational problem and adaptively extracts a finite number of band-limited modes, but its performance depends on the preset number of modes and penalty factors. SSA reconstructs the signal by performing eigenvalue decomposition on a trajectory matrix; however, its decomposition performance is highly sensitive to the choice of window length, which must be selected empirically [21]. As an adaptive decomposition technique, EMD does not require preset basis functions or parameters and can decompose complex signals into a series of intrinsic mode functions (IMFs), yet it suffers from mode mixing and endpoint effects, especially when the signal exhibits strong noise or closely spaced frequency components.

Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) improves upon EMD by introducing noise realizations that are matched to the characteristic scales of the signal components, thereby achieving a more complete and stable decomposition while effectively suppressing mode overlap. Specifically, EMD is applied to both the original signal and multiple noise-added realizations, and the IMFs of the same order are then ensemble-averaged to obtain the corresponding final modes. This ensemble strategy reduces the influence of random noise in any single realization and enhances the robustness of the

decomposition. The computational procedure can be summarized as follows:

(i) Obtain the first-order mode by adding white noise to the original signal, performing EMD on each noisy realization, and ensemble-averaging the first IMFs.

$$c_1(t)=\frac{1}{n}\sum_{i=1}^{n}\mathrm{IMF}_1\left(x(t)+\varepsilon w_i(t)\right) \tag{5}$$

where $\mathrm{IMF}_1(\cdot)$ denotes the first intrinsic mode function obtained by EMD, $w_i(t)$ is the added white-noise realization in the i-th ensemble, ε εis the noise amplitude coefficient, and n is the number of ensembles.

(ii) At the k-th modal extraction stage ($k \geq 2$), white noise is adaptively added to the residual signal $r_{k-1}(t)$ obtained from the previous iteration, and the corresponding mode is extracted in the form of a specific intrinsic component. The k-th intrinsic mode function $c_k(t)$ is obtained by ensemble-averaging the first IMFs of all noise-added realizations:

$$c_k(t)=\frac{1}{n}\sum_{i=1}^{n}\mathrm{IMF}_1\left(r_{k-1}(t)+\varepsilon w_{i,k}(t)\right) \tag{6}$$

where $w_{i,k}(t)$ is the k-th noise realization in the i-th ensemble.

(iii) After K iterations, the original signal $x(t)$ can be perfectly reconstructed as the sum of all intrinsic mode functions and the final residual term:

$$x(t)=\sum_{k=1}^{K}c_k(t)+r_K(t) \tag{7}$$

Where $c_k(t)$ denotes the k-th intrinsic mode function obtained by CEEMDAN and $r_K(t)$ is the final residual after decomposition.

VMD decomposes the input signal into a set of band-limited modes $u_k(t)$ by formulating the problem as a constrained variational optimisation. Each mode is assumed to be compact around its centre frequency in the spectral domain, and the optimal modes are obtained by minimising the sum of their bandwidths. The corresponding variational problem can be written as:

$$\min_{\{u_k\},\{\omega_k\}}\left\{\sum_{k=1}^{K}\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)\square u_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right\} \tag{8}$$

where K is the number of modes and ω_k is the center frequency of the k-th mode.

To ensure that the decomposed modes can fully reconstruct the original signal, VMD imposes a reconstruction constraint during optimisation, which requires that the sum of all modes is equal to the input signal $f(t)$. This constraint guarantees the invertibility and physical consistency of the decomposition, namely

$$\sum_{k=1}^{K}u_k(t)=f(t) \tag{9}$$

so that no information is lost and the original time series can be recovered exactly from the extracted modes.

3. Cleaning-Decomposition Synergistic Offshore Wind Power Forecasting Method

This section proposes a novel composite synergistic forecasting framework for offshore wind power based on adaptive optimization, aimed at addressing high data noise, strong non-stationarity, and the vulnerability of forecasting models to outliers. First, in view of the highly fluctuating operating data and numerous abnormal points in offshore wind farms, a two-stage cleaning strategy that combines DBSCAN clustering with ninth-order polynomial regression is developed to accurately identify and correct outliers in the power curve. Second, to enhance feature extraction from complex wind power series, a hybrid decomposition scheme, CEEMDAN-PSO-VMD, is constructed, in which PSO is used to adaptively optimize the VMD parameters and obtain more physically meaningful intrinsic mode components. Finally, a PSO-TCN-GRU hybrid forecasting model is designed. The TCN with dilated convolutions provides an extended receptive field for multi-scale feature learning, while the GRU captures temporal dependencies; PSO is further employed to optimize the model hyperparameters, thereby achieving high-precision offshore wind power forecasting.

3.1. Data Cleaning Methods

Data Sources and Characteristics Analysis

This study uses measured operational data from January 2022 to January 2023 for the 48 MW installed capacity at Site F1 of a municipal offshore wind farm for analysis, processing, and algorithm validation. The dataset consists of 37,163 SCADA records sampled at 15-minute intervals. It contains variables such as wind speed and active power output, and the original wind speed-power scatter plot is shown in Figure 2.

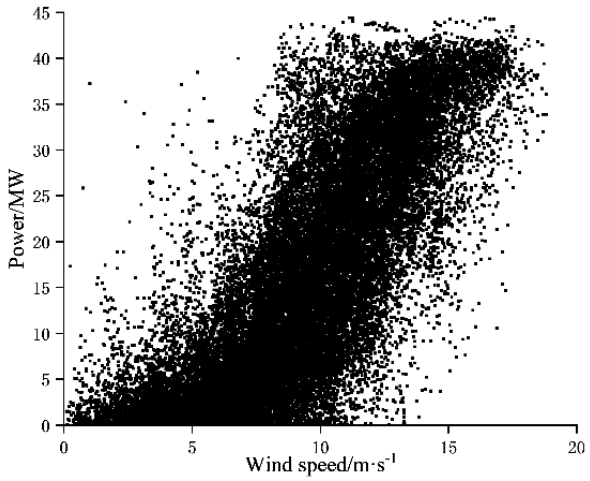

Figure 2. Original wind speed-power scatter plot

As shown in Figure 2, the overall trend indicates that power output increases with wind speed. However, the scatter of the sample points is relatively dispersed, and a number of points deviate markedly from the main trend, forming obvious outliers that cannot be explained by normal operating fluctuations. According to their distribution characteristics, these outliers can be divided into two categories:

(i) High-wind-speed, low-power points. For these points, the power output is significantly lower than that of most samples at the same wind speed, often forming an evident horizontal band or a downward-opening pattern in the scatter plot. This behaviour contradicts the physical power curve and usually reflects abnormal operating conditions of the turbines.

(ii) Extremely low-power points. These points exhibit very small power values, are sparsely distributed, and are clearly separated from the main data cluster. They typically correspond to shutdown, standby, or fault states that are not representative of normal power generation.

Such anomalies are mainly caused by unit maintenance, wind power curtailment, equipment failures, and communication delays. If they are not properly treated, they severely affect the fitting accuracy of the wind speed–power curve, distort the statistical characteristics of the dataset, and degrade the performance and stability of subsequent forecasting models. Therefore, the following data cleaning procedures focus on accurately identifying and treating these outliers so as to improve overall data quality and provide a more reliable basis for model training.

Data Cleaning Method Based on DBSCAN and Polynomial Regression

To address the complex distribution and diverse types of outliers in wind power data, this paper proposes a collaborative anomaly-cleaning method that combines DBSCAN clustering with ninth-order polynomial regression. Through a two-stage procedure of clustering-based identification and regression-based correction, the method enables accurate detection and effective repair of abnormal data points. Compared with traditional single-stage cleaning strategies, the proposed approach better preserves the integrity and continuity of the original time series under complex noise conditions. The overall workflow is illustrated in Figure 3.

This method adopts a two-stage collaborative cleaning procedure that targets both scattered outliers and clustered systematic anomalies in wind power data. First, the raw data are preprocessed to remove invalid records and perform normalization, thereby eliminating dimensional differences, improving the reliability of distance measures, and enhancing the stability of subsequent algorithms. In the first-stage cleaning module, DBSCAN is employed for density-based clustering. By exploiting local density patterns to distinguish normal operating points from sparse noise, DBSCAN provides strong detection capability for random, isolated anomalies that frequently occur in wind power datasets. After this step, the processed data are denormalized to restore their actual physical scale. In the second-stage cleaning module, a ninth-order polynomial regression model is used to fit the wind speed-power relationship and to identify and correct structurally clustered anomalies. The goodness of fit is evaluated using the coefficient of determination, ensuring that the corrected data are consistent with the true trend of the power curve. Overall, the proposed cleaning strategy effectively suppresses local anomalous disturbances while preserving the global structural characteristics of the sequence, thereby providing a more reliable data foundation for subsequent forecasting models.

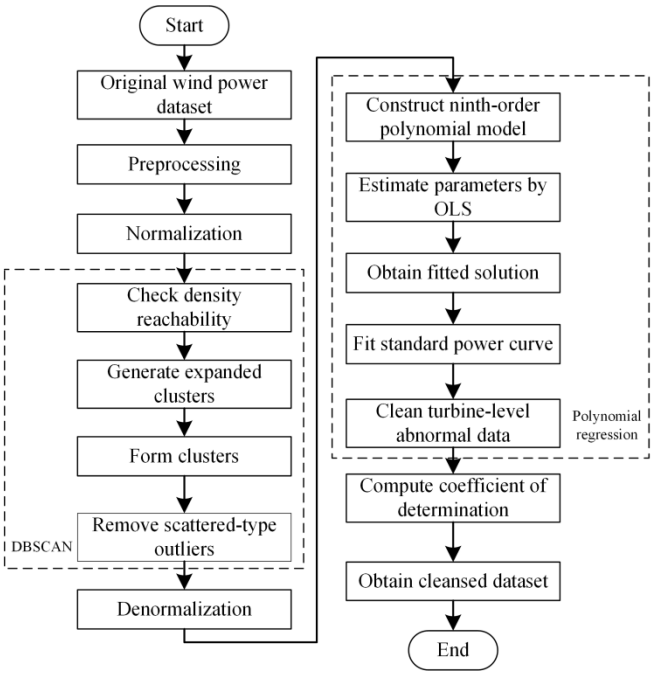

Figure 3. Two-stage data cleaning flowchart

3.2. Offshore Wind Power Forecasting Algorithm

The offshore wind power forecasting framework proposed in this paper is shown in Figure 4. It consists of three main components: a PSO-based optimization module, a two-stage decomposition module, and a TCN-GRU prediction module.

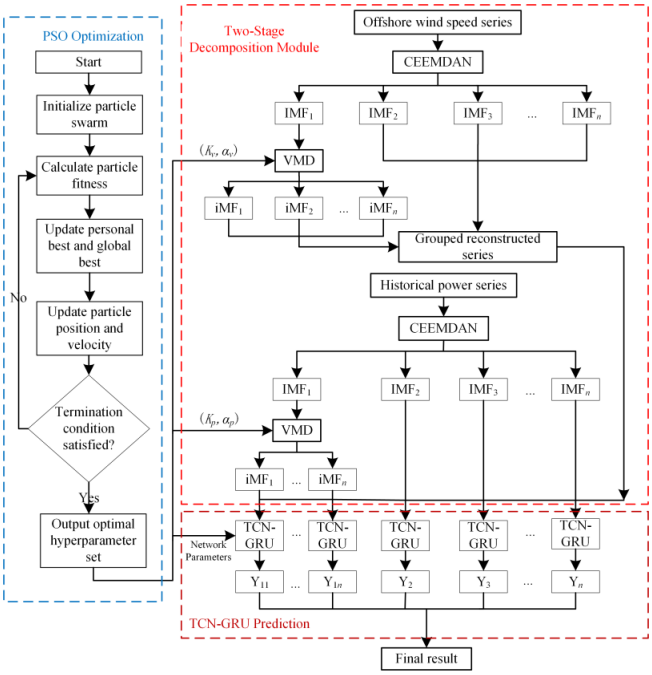

Figure 4. Offshore wind power forecasting model

PSO Optimization Module

This module provides optimal parameter configurations for the subsequent VMD decomposition and TCN-GRU prediction stages. In conventional approaches, key parameters such as the number of modes, penalty factors, and network hyperparameters are typically selected empirically, which can easily lead to under- or over-decomposition, as well as mismatches between model structure and data characteristics, thereby constraining forecasting accuracy and robustness. To overcome these limitations, this paper introduces a particle swarm optimization (PSO) algorithm. Using sample entropy as the fitness function, PSO automatically searches for key parameters, including the VMD mode number K and penalty factor α, as well as the convolution kernel size, number of filters, and GRU hidden-layer dimension in the TCN-GRU network. In this way, both the two-stage decomposition module and the prediction module operate under near-optimal parameter settings.

Two-Stage Decomposition Module

This module is designed to handle the strong non-stationarity and complex noise components in offshore wind speed and power sequences, for which single-stage decomposition methods often suffer from mode aliasing. It performs multiscale refined decomposition and feature reconstruction of the original signals. First, CEEMDAN is applied to the cleaned wind speed and historical power series, decomposing the original signals into multiple IMFs ordered from high to low frequency. This step partially mitigates non-stationarity and separates components at different scales. However, the high-frequency components, especially IMF1, still contain substantial noise and overlapping frequency bands, which may degrade forecasting performance if used directly for modeling. Therefore, under PSO-optimized parameters, IMF1 is subjected to a secondary decomposition using VMD. The resulting sub-modes are then regrouped and reconstructed with the remaining IMFs according to their feature similarity. This process yields wind speed and power components with clearer multiscale structures and reduced noise interference. The refined components provide smoother and more discriminative input features for the prediction module, thereby enhancing the model's ability to extract key information and improving forecasting accuracy.

TCN-GRU Prediction Module

This module takes the reconstructed components from the two-stage decomposition module as inputs to perform multi-scale power forecasting. Compared with traditional architectures such as RNN and LSTM, a single model often struggles to capture both rapid local fluctuations and long-term trends simultaneously, which can result in underfitting or excessive smoothing. To address this issue, a hybrid TCN-GRU architecture is adopted. The TCN employs dilated convolutions to achieve a long effective receptive field,

enabling the extraction of multi-scale local temporal features, while the GRU captures long-term dependencies and nonlinear dynamics within the sequences. On this basis, PSO is further used to perform global optimization of key hyperparameters in the prediction model, thereby enhancing overall forecasting accuracy and robustness. Finally, the forecasting results obtained for each decomposed component are aggregated to produce the final offshore wind power prediction.

4. Experimental Results

4.1. Data Cleaning Results

The original offshore wind power dataset was cleaned using the proposed method. Figure 5 shows the results after DBSCAN clustering. The silhouette coefficient S was adopted as the evaluation metric for clustering performance, where a larger S indicates better clustering quality. After multiple iterations of parameter tuning, the algorithm parameters were set to $\varepsilon = 0.02$ and $\mathrm{M} = 5$, corresponding to the maximum silhouette coefficient. As shown in Figure 5, the scattered anomalous points are effectively removed, although some clustered anomalies still remain.

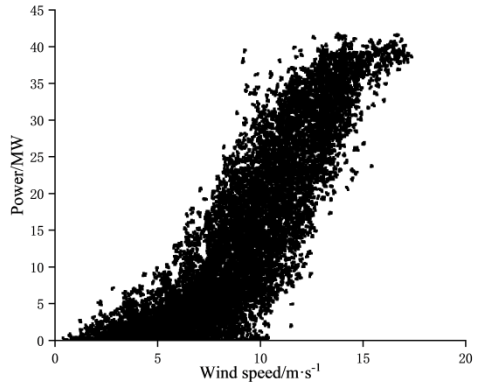

Figure 5. DBSCAN-based cleaning results

On the basis of the DBSCAN clustering, a ninth-order polynomial function was then fitted to the data. Using the 3σ criterion to identify outliers, an updated dataset with anomalous points removed was obtained, yielding a coefficient of determination of $R^2 = 0.8695$. The corresponding polynomial fitting result is presented in Figure 6.

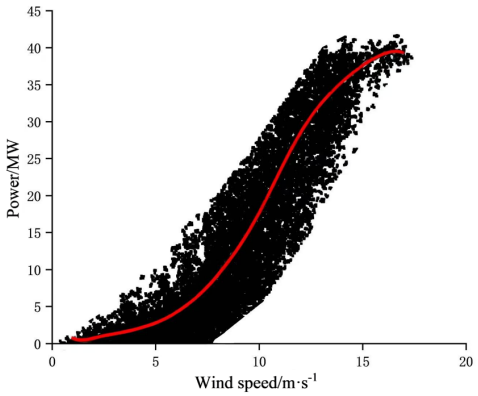

Figure 6. Polynomial fitting results

During the removal of abnormal data, missing values may be introduced into the original dataset. To ensure the temporal completeness of the wind power time series, these missing values must be imputed. In this study, a linear interpolation scheme is adopted: a fitting function is constructed using the data points immediately before and after the missing value, and the corresponding value is then used to fill the gap. The interpolation formula is given by:

$$\varphi(x) = \frac{(x - x_2)}{(x_1 - x_2)} y_1 + \frac{(x - x_1)}{(x_2 - x_1)} y_2 \tag{10}$$

where (x_1, y_1) and (x_2, y_2) denote the data points preceding and following the missing value, respectively.

The Spearman correlation coefficient was employed to compare the proposed cleaning method with three alternative approaches, namely DBSCAN, polynomial regression, and polynomial regression combined with DBSCAN, as well as with the uncleaned data. The results are summarized in Table 1.

Table 1. Spearman correlation after data cleaning

Cleaning method	Spearman
Original data	0.7977
DBSCAN	0.8514
Polynomial regression	0.8529
Polynomial regression + DBSCAN	0.8893
DBSCAN+ Polynomial regression	0.9062

As shown, all cleaning strategies increase the Spearman correlation between wind speed and power relative to the original dataset (0.7977), indicating that data preprocessing is necessary. DBSCAN and polynomial regression alone improve the correlation to 0.8514 and 0.8529, respectively,

corresponding to an absolute increase of about 0.05. When the two methods are combined in the order of polynomial regression followed by DBSCAN, the correlation is further enhanced to 0.8893. In contrast, the proposed DBSCAN + polynomial regression method achieves the highest correlation of 0.9062, representing an absolute improvement of 0.1085 over the raw data (approximately a 13.6% relative increase) and about 0.055 over using either DBSCAN or polynomial regression alone. This demonstrates that the collaborative strategy of first removing scattered anomalies with DBSCAN and then correcting clustered deviations via polynomial regression provides the greatest enhancement in data consistency and yields the most reliable cleaned dataset for subsequent forecasting.

4.2. Wind Speed-Power Sequence Decomposition Results

After data cleaning and interpolation, the offshore wind speed series was first decomposed by CEEMDAN into 10 intrinsic mode functions (IMFs). IMF1, identified as the key high-frequency component, was then further decomposed by VMD into 5 sub-modes, while IMF2–IMF10 were retained unchanged. Following feature-based regrouping and reconstruction, a total of 14 wind speed components were obtained. Similarly, the offshore wind power series was decomposed by CEEMDAN into 10 IMFs, and the corresponding IMF1 was further decomposed by VMD into 4 sub-modes. After recombination, 13 wind power components were derived.

Taking the offshore wind speed signal as an example, the CEEMDAN decomposition results are shown in Figure 7 (only the first five components are plotted). It can be seen that the original signal exhibits pronounced abrupt fluctuations with no clear periodic pattern. After CEEMDAN decomposition, the nonlinearity and randomness of the signal are substantially reduced, and the volatility of the original series is effectively suppressed, with different IMFs capturing distinct time-scale characteristics.

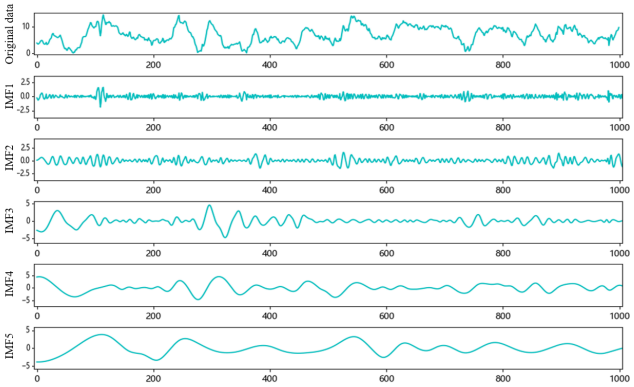

Figure 7. CEEMDAN decomposition of the wind speed sequence (first five components)

On this basis, the PSO-VMD method is applied to further decompose IMF1, producing a set of more stable and refined high-frequency sub-series, as illustrated in Figure 8. The VMD sub-modes show relatively narrow-band oscillations with reduced noise and weaker mutual interference, indicating that the secondary decomposition successfully isolates useful high-frequency structures from random disturbances.

Overall, the combined CEEMDAN and PSO-VMD decomposition strategy provides a clear multi-scale representation of the offshore wind speed and power sequences. Low-frequency components describe the underlying trend and periodic variations, while high-frequency components capture local fluctuations and transient disturbances. This hierarchical decomposition effectively mitigates mode mixing, enhances the separability of features at different time scales, and lays a solid foundation for the subsequent TCN-GRU prediction model to learn both global trends and fine-grained dynamics more accurately.

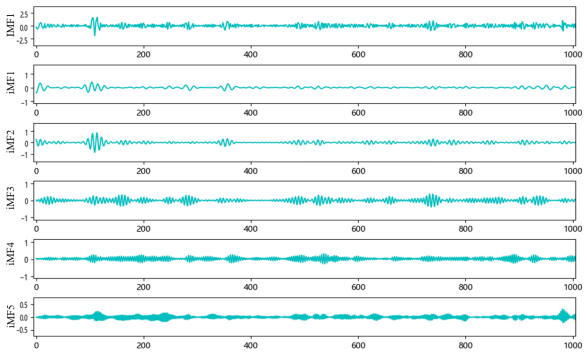

Figure 8. VMD decomposition results for the IMF1 component

4.3. Power Prediction Results

In this study, Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) are adopted as performance metrics. Smaller values of these indicators correspond to lower forecasting errors and hence higher prediction accuracy. The definitions and calculation formulas of these metrics are summarized in Table 2.

Table 2. Main Evaluation Metrics

Evaluation metric	Calculation formula
MSE	$\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$
RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$

MAE	$\frac{1}{n}\sum_{i=1}^{n}\lvert\hat{y}_i - y_i\rvert$
MAPE	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{\hat{y}_i - y_i}{y_i}\right\rvert$

Comparison with Typical Models

To evaluate the predictive performance of the proposed model, three representative benchmark models are selected for comparison, namely the Transformer, LSTM, and GRU networks. These models are widely used in time-series forecasting and provide a reasonable basis for assessing the advantages of the proposed approach. As shown in Table 3 and Figure 9, there are pronounced differences among the models across all evaluation metrics. The Transformer model shows limited capability in capturing local temporal features, yielding an MSE of 4.36 MW, which corresponds to the highest overall error level. LSTM and GRU improve the modeling of temporal dependencies, reducing the MSE to 3.38 MW and 3.34 MW, respectively; however, their accuracy gains remain constrained due to the lack of multi-scale feature decomposition and dedicated anomaly handling. Taking the GRU model as a baseline, the proposed method further reduces MSE, RMSE, MAE, and MAPE by 2.59 MW, 2.54 MW, 1.51 MW, and 3.41%, respectively, corresponding to an error reduction of approximately 57%-78%. Compared to conventional single-model approaches, the proposed methodology demonstrates distinct advantages in prediction accuracy and operational stability. This improvement is achieved through an integrated framework that systematically addresses key challenges in wind power forecasting: a two-stage data cleaning process enhances input quality, secondary decomposition effectively handles signal non-stationarity, and the PSO-optimized TCN-GRU hybrid architecture successfully captures both long-range dependencies and complex temporal dynamics. The synergistic combination of these components enables more robust and reliable power output predictions under diverse operating conditions.

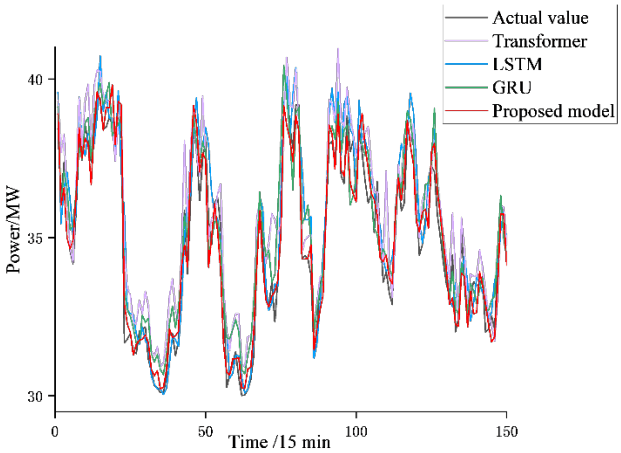

Figure 9. Comparison of power prediction results obtained by different method

Table 3. Performance comparison with classical models

Model	Evaluation metrics			
	MSE/MW	RMSE/MW	MAE/MW	MAPE/%
Transformer	4.36	3.74	3.49	7.39
LSTM	3.38	3.49	2.25	6.12
GRU	3.34	3.41	2.14	5.93
Proposed model	0.75	0.87	0.63	2.52

Comparison of Decomposition Strategies

Building on the comparison with the three typical models and taking the TCN-GRU model as a reference, this study further investigates the effects of two-stage decomposition and wind speed–power coupling information on forecasting performance. With the network architecture kept unchanged, the composition of the input sequences and the decomposition strategies are modified to construct three comparative models:

(i) Model 1: Two-stage decomposition applied only to the offshore wind speed series;

(ii) Model 2: Two-stage decomposition applied only to the offshore wind power series;

(iii) Model 3: Two-stage decomposition applied simultaneously to both the wind speed and wind power series.

Table 4. Prediction error comparison for different models

Model	Evaluation metrics			
	MSE/MW	RMSE/MW	MAE/MW	MAPE/%
TCN-GRU	3.09	3.02	3.02	3.02
Model 1	1.97	1.96	1.96	1.96
Model 2	2.06	2.08	2.08	2.08
Model 3	1.22	1.46	1.46	1.46
Proposed method	0.75	0.87	0.87	0.87

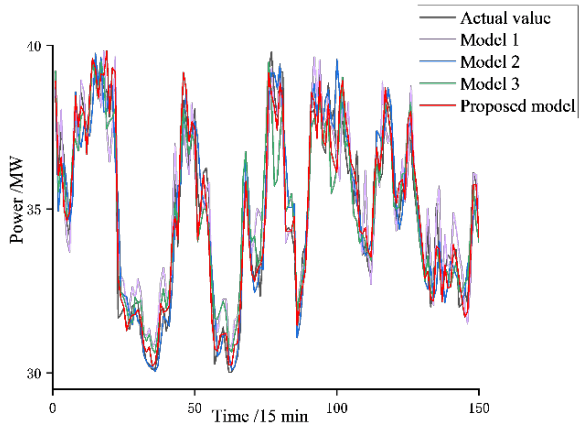

Figure 10. Comparison of Prediction Results for Different Models

As shown in Table 4 and Figure 10, comparing Models 1 and 2 with the baseline TCN-GRU model indicates that applying two-stage decomposition to either the wind speed series or the historical power series improves forecasting performance. Relative to direct prediction by TCN-GRU, Model 1 reduces MSE, RMSE, MAE, and MAPE by 1.12 MW, 1.06 MW, 0.45 MW, and 1.23%, respectively, while Model 2 achieves corresponding reductions of 1.03 MW, 0.94 MW, 0.41 MW, and 1.08%. Model 3, which performs two-stage decomposition on both wind speed and power series, further decreases these four error indices by 0.84 MW, 0.62 MW, 0.57 MW, and 0.96% compared with Model 2. In turn, the proposed method yields additional reductions of 0.47 MW, 0.59 MW, 0.41 MW, and 0.86% relative to Model 3. These results demonstrate that incorporating two-stage decomposition and optimizing the TCN-GRU hyperparameters via PSO can significantly enhance the predictive performance of the model.

5. Conclusion

To mitigate the loss of forecasting accuracy caused by poor data quality and non-stationary time series in offshore wind power applications, this paper proposes a novel composite collaborative forecasting method based on intelligent optimization. Case study simulations lead to the following conclusions:

(i) The proposed two-stage data cleaning algorithm, which combines DBSCAN with ninth-order polynomial regression, can simultaneously identify both scattered and clustered outliers, thereby improving the quality and consistency of the input data;

(ii) By introducing PSO to jointly and adaptively optimize the VMD decomposition parameters and the TCN-GRU network hyperparameters, the uncertainty associated with empirical parameter tuning is reduced, and the robustness and adaptability of the forecasting model are significantly enhanced;

(iii) On the basis of data cleaning, the coordinated operation of the two-stage wind speed-power decomposition module and the prediction module forms an integrated cleaning-decomposition-prediction framework, which effectively improves the accuracy and stability of offshore wind power forecasting.

The model developed in this study currently focuses on wind speed as the core variable to ensure framework robustness. Incorporating additional meteorological variables such as wind direction and sea surface temperature is expected to enhance prediction accuracy, particularly improving model robustness and adaptability under extreme weather conditions. Wind direction directly impacts the efficiency of wind turbines facing the wind, while sea surface temperature influences wind field characteristics by modulating air-sea interactions. Therefore, future work will validate the model's performance under extreme weather conditions and explore integrating these multi-physics variables to develop a more precise and robust next-generation prediction system.

Acknowledgements.

This research was supported by the Science and Technology Project of State Grid Shanghai Electric Power Company (No.52090R250007): Research on Multi-Objective Optimization Control and Risk Prevention Technology for Far-reaching Offshore Wind Turbines and Clusters Driven by Multi-Source Heterogeneous Data.

References

[1] Zhang XH, Zhang JF, He YG, et al. Imperfect maintenance decision for wind turbines based on multi-state space partitioning. Acta Energiae Solaris Sinica. 2022; 43(11): 203-214.

[2] Zhao L, Wei C, Wang Y, et al. Macrositing of offshore wind farms and estimation of wind energy reserves. Acta Energiae Solaris Sinica. 2024; 45(5): 1-8.

[3] Ge C, Yan J, Liu YQ, et al. Review of key technologies for operation, control and maintenance of offshore wind farms. Proceedings of the CSEE. 2022; 42(12): 4278-4292.

[4] WANG G, JIA R, LIU J H, et al. A hybrid wind power forecasting approach based on Bayesian model averaging and ensemble learning[J]. Renewable Energy, 2020; 145:2426-2434.

[5] AMBACH D, SCHMID W. A new high-dimensional time series approach for wind speed, wind direction and air pressure forecasting[J]. Energy, 2017; 135:833-850.

[6] ZHOU YY, MA LW, NI WD, et al. Data enrichment as a method of data preprocessing to enhance short-term wind power forecasting[J]. Energies, 2023; 16(5): 2094.

[7] Zhao YN, Ye L, Zhu QW. Characteristics and processing method of curtailment-related abnormal data clusters in wind farms. Automation of Electric Power Systems. 2014; 38(21): 39-46.

[8] Zhu QW, Ye L, Zhao YN, et al. Identification and reconstruction method for abnormal output power data of wind farms. Power System Protection and Control. 2015; 43(3): 38-45.
[9] Zhao WQ, Wang JF, Guo DQ. Health status assessment of wind turbines based on ConvLSTM-SA. Electric Power Information and Communication Technology. 2022; 20(11): 20-26.
[10] Mo FY, Wang WH, Guo Q. Abnormal wind power data processing method for wind turbines based on MADM-QM. Renewable Energy. 2025; 43(3): 339-345.
[11] ZHAO X Y, LIU J F, YU D R, et al. One-day-ahead probabilistic wind speed forecast based on optimized numerical weather prediction data[J]. Energy Conversion and Management, 2018; 164: 560-569.
[12] SUN Z X, ZHAO M Y. Short-term wind power forecasting based on VMD decomposition, ConvLSTM networks and error analysis[J]. IEEE Access, 2020; 8: 134422-134434.
[13] Qi CC, Wang XW. Short-term offshore wind power prediction considering wind direction and atmospheric stability. Power System Technology. 2021; 45(7): 2773-2780.
[14] Xie JC, Yu QC, Wang ZC, et al. Short-term wind power forecasting based on VMD-GRAU. Guangxi Sciences. 2024; 31(4): 773-780,787.
[15] Sun S, Wei L, Xu J, Jin Z. A New Wind Speed Forecasting Modeling Strategy Using Two-Stage Decomposition, Feature Selection and DAWNN. Energies. 2019; 12(3):334.
[16] Li A, Ran HJ, Li LW, et al. Short-term wind power forecasting based on CEEMDAN-TCN. Modern Electronics Technique. 2025; 48(2): 97-02.
[17] WANG G, JIA R, LIU J H, et al. A hybrid wind power forecasting approach based on Bayesian model averaging and ensemble learning[J]. Renewable Energy, 2020, 145:2426-2434.
[18] Ma LY, Sun JM, Yu SL, et al. Abnormal operating condition early warning for wind turbines based on DBSCAN and SDAE. Journal of Chinese Society of Power Engineering. 2021; 41(9): 786-793,808.
[19] Shao L, Huang W, Liu H, et al. Study of Wind Power Prediction in ELM Based on Improved SSA[J]. IEEJ Transactions on Electrical and Electronic Engineering, 2025; 20(6): 853-861.
[20] Yang M, Chen X, Du J, et al. Ultra-Short-Term Multistep Wind Power Prediction Based on Improved EMD and Reconstruction Method Using Run-Length Analysis[J]. IEEE Access, 2018; 6: 31908-31917.
[21] ZHANG X Y, LI C S, WANG X B, et al. A novel fault diagnosis procedure based on improved symplectic geometry mode decomposition and optimized SVM[J]. Measurement, 2021; 173: 108644.