Journal of Electronic Research and Application, 2026, Volume 10, Issue 8
http://ojs.bbwpublisher.com/index.php/JERA
Online ISSN: 2208-3510
Print ISSN: 2208-3502

# Research on Medium- and Long-Term Wind Power Forecasting Technology Based on Ensemble Forecasting

**Li Liu, Xiaolin Xu, Qian Liang, Jiayu Chen, Sen Duan**

State Power Investment Corporation Energy Science and Technology Research Institute, Shanghai 201109, China



**Abstract:** With the continuous growth of grid-connected wind power capacity and the full participation of new energy in power market transactions, medium and long-term power forecasting has become a key supporting technology to guarantee the safe and economic operation of power systems. The forecasting ability of deterministic numerical weather prediction (NWP) decays sharply for lead times longer than 4 days. Ensemble forecasting provides a probabilistic information foundation to break this bottleneck, yet its application at wind farm scale is confronted with critical problems including coarse spatial resolution, prominent systematic bias, and low utilization rate of multi-member information. Taking Yuanqu Wind Farm in Shanxi Province as the research object, this paper systematically carries out research on multi-source ensemble forecast data adaptation processing, hybrid forecasting modeling driven by random forest and physical models, as well as probabilistic forecasting methods focusing on medium- and long-term wind power forecasting based on ensemble forecasts. In terms of data adaptation processing, four global ensemble forecasting systems including ECMWF-ENS, UKMO-EGRR, CMA-BABJ and NCEP-GEFS are compared and evaluated for their applicability in complex mountain wind farms. The results show that ECMWF-ENS has the lowest forecasting error and optimal seasonal stability. On this basis, a complete processing workflow consisting of spatial downscaling, quantile mapping bias correction, and ensemble spread consistency reconstruction is established, which reduces the root mean square error (RMSE) of wind speed forecasting by 15–25%. In terms of forecasting modeling, a hybrid driving framework combining random forest wind speed correction and equivalent power curve physical conversion is proposed. The random forest is adopted to learn local nonlinear residuals, while physical models provide hard constraints for energy conversion. The model is trained with data from January 2024 to February 2025 and verified in the independent test period from January to April 2026. The power generation accuracy of ensemble mean point forecasting reaches 83.26%, with an RMSE of 42.73 MW, meeting the requirements of power industry standards. This study constructs a complete technical chain from raw ensemble forecast data to wind farm-level probabilistic power forecasting, and verifies the effectiveness and engineering practicability of the hybrid driving method for medium and long lead times.

**Keywords:** Medium and long-term wind power forecasting; Ensemble forecasting; Random forest; Physical-data hybrid driving

**Online publication:** Sept 4, 2026

## 1. Introduction

Driven by global energy transition and carbon neutrality goals, wind power and photovoltaic power are accelerating to become the mainstay of power supply [1]. By the end of 2025, China's installed wind power capacity reached 640 million kilowatts, and the cumulative installed proportion of wind and photovoltaic power hit 47.3%, exceeding thermal power for the first time in history. The strong randomness, intermittency, and fluctuation of high-proportion new energy transform power systems from the traditional "load-following generation" mode to the collaborative interaction mode of "source-grid-load-storage", putting forward higher requirements for the lead time and accuracy of power forecasting [2,3]. Meanwhile, the Circular of the National Development and Reform Commission on Deepening the Market-oriented Reform of New Energy On-grid Electricity Prices and Promoting High-quality Development of New Energy (Document No. FGJG (2025) 136) clearly promotes the full access of new energy to the power market, and power transactions urgently require the support of high-precision meteorology and power forecasting technologies.

Medium and long-term power forecasting (4–30 days in advance) serves as core information supporting power balance, maintenance scheduling, cross-provincial and cross-regional transactions, and reserve optimization [4–6]. At present, forecasting technologies relying on deterministic numerical weather prediction are only effective within 0–3 days. After the 4th day, the nonlinear chaotic effect of the atmosphere leads to a sharp rise in errors of single deterministic forecasts [7–10]. By constructing multiple groups of initial value and physical perturbation members to describe the probability density distribution of forecasting states, ensemble forecasting acts as a fundamental approach to break the lead time bottleneck. Taking the ECMWF global ensemble forecasting system as an example, it provides 51-member forecasts for 0–15 days and subseasonal forecasts extended to 46 days, offering abundant probabilistic information for wind and solar power forecasting for 4–30 days [11,12].

Nevertheless, ensemble forecasting has relatively coarse spatial resolution (about 18–36 km), resulting in scale mismatch with actual meteorological exposure of wind farms; each member suffers from systematic bias and inaccurate ensemble spread [13]. Therefore, research on multi-timescale wind power forecasting based on ensemble forecasting is of great significance to break the bottleneck of medium and long-term new power cognition for power grids and improve the safety and economy of high-proportion new energy systems.

At present, ECMWF Ensemble Forecast (ENS), ECMWF Subseasonal Forecast, and the US NCEP Global Ensemble Forecast System (GEFS) are widely applied in renewable energy forecasting businesses of many countries. Relevant studies generally prove that ensemble mean forecasting can effectively smooth random errors caused by internal atmospheric variability on the 4–10 day scale and outperform single deterministic forecasts. However, the RMSE of ensemble mean forecasting still increases significantly for lead times over 10 days. Domestic scholars have carried out preliminary adaptive research on ensemble forecasting in medium- and long-term wind/photovoltaic forecasting targeting China's complex terrain and climate characteristics, yet obvious deficiencies exist in refined preprocessing at wind farm scale and deep coupling with power models [14–17].

## 2. Technical route

Aiming at the above problems and challenges, this project centers on the main line of "data foundation–forecasting core–decision support" and sets three progressive research contents:

First, characteristic analysis and adaptive processing of multi-source ensemble forecast data. This step systematically analyzes the uncertainty features of typical ensemble forecast systems, investigates spatial downscaling methods at the wind farm station scale, and develops quantile mapping bias correction as well as ensemble spread consistency reconstruction technologies. Second, construction of wind power prediction models based on ensemble forecasts. Targeting two time scales of 4–15 days and 15–30 days, a hybrid driving model combining random forest-corrected wind speed and physical power curves is established. Third, system integration and engineering verification. The above methods are integrated into a station-level prediction system, whose prediction accuracy and practical adaptability are verified using actual operational data.

This study takes the Yuanqu Wind Farm in Shanxi Province as a typical research object. The prediction horizons are set as 10 days and 20 days with a 15-minute temporal resolution, plus a 30-day horizon adopting daily average power, so as to meet the demands of medium and long-term power trading. The training and validation period spans from January 2024 to December 2025, and the independent testing period covers January to April 2026, with actual grid-connected power from SCADA data adopted as the verification benchmark.

## 3. Characteristic analysis and adaptive processing of multi-source ensemble forecast data

The performance of medium and long-term wind power forecasting highly depends on the quality and adaptability of input meteorological forecast data. Starting with characteristic analysis of multi-source ensemble forecast data, this chapter compares and verifies the superiority of ECMWF as the core data source. Targeting the demands of wind power generation, it systematically carries out meteorological element extraction and quality control, spatial downscaling and site adaptation, multi-timescale feature construction, as well as bias correction and consistency reconstruction of ensemble members, laying a solid data foundation for subsequent forecasting modeling.

### 3.1. Characteristic analysis of multi-source ensemble forecast data

Based on wind speed forecast verification data of Yuanqu Station in Shanxi Province from January to November 2024, comprehensive analysis is conducted on the data quality of four global ensemble forecasting systems, including ECMWF-ENS (ECMF), UKMO-EGRR (EGRR), CMA-BABJ (BABJ), and NCEP-GEFS (KWBC). Judging from monthly overall statistical indicators RMSE and MAE, ECMWF always achieves the optimal performance, with monthly average RMSE fluctuating between 0.90 m/s (April) and 1.24 m/s (March), and corresponding MAE ranging from 0.78 m/s (October) to 0.90 m/s (March), featuring the lowest error and favorable seasonal stability. EGRR ranks second, with monthly RMSE from 1.09 m/s (April) to 1.40 m/s (March) and MAE from 0.83 m/s (October) to 1.05 m/s, also showing high forecasting reliability. BABJ has significantly higher errors than the former two, with monthly RMSE between 1.32 m/s (October) and 1.74 m/s (March) and MAE from 1.04 m/s (October) to 1.47 m/s, accompanied by drastic seasonal fluctuations, which indicates systematic deficiencies in physical processes or initial field processing of the model under complex terrain. KWBC delivers the worst performance in all months, with the minimum monthly RMSE of 1.46 m/s (April) and maximum of 2.31 m/s (January), and MAE ranging from 1.46 m/s to 1.79 m/s. Its errors rise sharply in high-winter periods, revealing severe inadaptability to mountain wind farm forecasting.

From the perspective of time series variation, errors of all systems are higher in winter (January, March), decrease in spring (April, May), stay low in summer (July, August), and rise slightly in autumn (September–November). This rule is related to strong northwest winds and atmospheric stability changes in the Yuanqu area, and also reflects the greater forecasting difficulty of numerical models for high wind speed events. It is noteworthy that ECMF and EGRR maintain a low error growth slope in all months throughout the year, while errors of BABJ and KWBC accelerate with the extension of lead time, especially KWBC. Based on 11-month data evaluation, ECMF and EGRR can be adopted as core input sources for wind power forecasting systems; BABJ requires bias correction before application, and KWBC cannot be directly used for wind speed forecasting and must undergo intensive correction or downscaling processing.

### 3.2. Bias correction and consistency reconstruction of ensemble members

According to wind speed forecast verification results of Yuanqu Station from January to November 2024, different global ensemble forecasting systems exhibit prominent systematic bias and error growth characteristics. Among all systems, ECMWF-ENS (ECMF) has the minimum RMSE and MAE across all lead times and months with optimal seasonal stability, so subsequent wind power forecasting research mainly relies on ECMF data. Nevertheless, raw ECMF outputs are global grid fields with a spatial resolution of about 9 km, which cannot directly characterize local wind features of mountainous terrain where Yuanqu Station is located. Even after spatial interpolation, accumulated systematic biases still exist with the extension of lead times. To improve the quantitative accuracy and probabilistic reliability of ECMF ensemble forecasting for local wind power prediction, spatial downscaling and site matching, systematic bias correction and ensemble spread consistency reconstruction must be implemented in sequence. This section follows the technical route of "from space to numerical value, from mean value to distribution".

ECMF ensemble forecasting contains multiple perturbation members, and its theoretical spread should reflect forecasting uncertainty. Verification shows that the ensemble spread of ECMF at Yuanqu Station is often smaller than actual errors (under-dispersion), especially for short lead times. This leads to excessively narrow confidence intervals of probabilistic forecasting and reduces the reliability of ensemble forecasting. Therefore, a method combining variance inflation and member reordering is adopted for consistency reconstruction. It is expected that after spatial downscaling and statistical correction, the RMSE of ECMWF wind speed forecasting can be reduced by 15–25%, and the reliability of probabilistic forecasting will be significantly improved. Corrected ECMF ensemble members will be taken as inputs for subsequent wind power forecasting models to enhance the accuracy of wind power curve conversion and the confidence level of probabilistic forecasting.

In summary, spatial downscaling, site matching, bias correction, and consistency reconstruction implemented on ECMF can acquire more accurate and reliable wind speed probabilistic forecasts, providing high-quality inputs for the wind power forecasting model in the next chapter.

## 4. Physical-data hybrid driven wind power forecasting model based on ensemble forecasting

### 4.1. Overview

Aiming at the problems of fast error accumulation of single physical models and poor physical consistency of pure data models, a hybrid driving architecture integrating random forest wind speed correction and physical power curve conversion is proposed. The random forest learns local nonlinear residuals of wind speed forecasts, while physical models provide hard constraints for energy conversion, achieving complementary advantages of the two methods.

## 4.2. Physical modeling and baseline construction of wind power

### 4.2.1. Physical model of wind turbine power curve

The theoretical relationship between wind turbine output power and hub-height wind speed is determined by aerodynamics. For wind speeds below rated speed, power is approximately cubic to wind speed:

$$P_{ideal}\left(v\right)=\begin{cases} 0v<v_{cut-in} \\ \dfrac{1}{2}\rho AC_p\left(\lambda,\beta\right)v^3 v_{cut-in}\le v<v_{rated} \\ P_{rated}v_{rated}\le v<v_{cut-out} \\ 0v\ge v_{cut-out} \end{cases}$$

Where: $\rho$ stands for air density ($\text{kg/m}^3$), A is the swept area of the rotor ($\text{m}^2$), $C_p$ is the wind power utilization coefficient (theoretical Betz limit 0.593, actual value 0.4–0.5), $\lambda$ is the tip speed ratio, and $\beta$ is the pitch angle. In actual operation, measured power curves deviate from theoretical curves due to turbulence, wake effect, blade fouling, and control strategies. Therefore, equivalent power curves based on wind farm SCADA data are established as physical baselines.

### 4.2.2. Equivalent power curve fitting based on binning regression

Wind turbine theoretical power is determined by aerodynamics, while actual curves deviate due to turbulence, wake, and blade contamination. Based on SCADA data of Yuanqu Wind Farm (single unit capacity 2.5 MW, hub height 90 m), equal-width binning (0.5 m/s interval) is adopted to construct equivalent power curves. Abnormal operating conditions are eliminated, and the median power of each wind speed bin is calculated. Local weighted scatterplot smoothing (LOWESS) is applied to smooth bin medians to eliminate random noise and obtain the final physical baseline. , as shown by the red line in **Figure 1**.

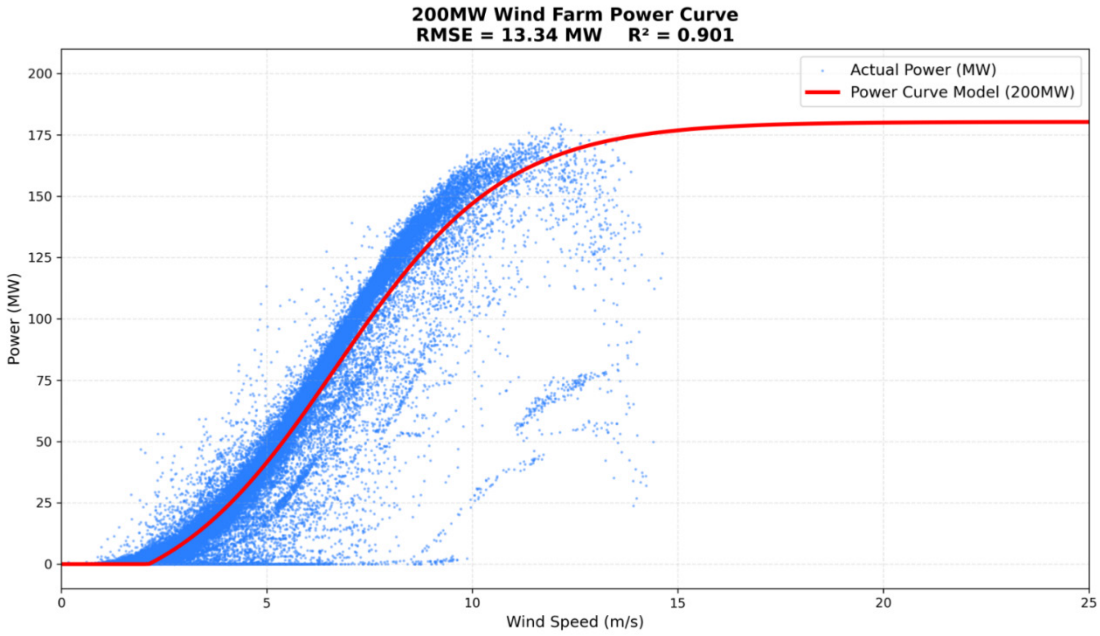


**Figure 1.** Equivalent power curve of Yuanqu wind farm.

## 4.3. Wind speed correction based on random forest

### 4.3.1. Characteristic analysis of wind speed forecasting errors

The bias correction in Chapter 2 mainly eliminates overall systematic shifts, while ECMF outputs still contain non-stationary residuals varying with meteorological conditions. Analysis of historical data from January to November 2024 shows that wind speed forecasting error $e=V_{ECMF}-V_{obs}$ has significant correlation with the following factors:

First, wind speed magnitude: ECMWF systematically overestimates wind speed in low-wind-speed areas (< 4 m/s) with a mean bias of + 0.6 m/s, while it underestimates wind speed in high-wind-speed areas (> 12 m/s) with a mean bias of -0.8 m/s. Second, atmospheric stability: the forecast error of near-surface wind speed increases significantly under stable nocturnal boundary layer conditions. Third, ensemble spread: when the spread among ensemble members is large (> 1.5 m/s), the forecast features high uncertainty and a larger absolute error. Fourth, topographic complexity: wind speed errors show opposite signs in valley and ridge locations, with underestimation in valleys and overestimation on ridges.

Simple overall bias correction cannot eliminate local nonlinear errors, so data-driven methods are introduced to learn complex error mapping relationships.

### 4.3.2. Design of random forest regression model

Random forest (RF) is an ensemble learning algorithm that reduces variance by constructing multiple decision trees and averaging prediction results. Compared with deep neural networks, RF is robust to high-dimensional features and missing values, capable of outputting feature importance, less prone to overfitting, and fast in training.

(1) Input feature design

An independent RF model is built for each ensemble member. $m$ (51 parallel training models). Input features $X_{m,t}$ fall into four categories (see **Table 1**):

| Category | Specific features | Physical significance |
|---|---|---|
| Forecast wind speed | $V_{m,t}, V_{mean,t}, V_{max,t}, V_{min,t}$ | Information of single member and ensemble statistics |
| Ensemble uncertainty | Ensemble standard deviation, inter-quantile range ($Q_{90}$-$Q_{10}$) | Indicator of forecasting credibility |
| Atmospheric environment | Wind direction, air density, geopotential height, time coding (month/day/hour) | Characterize local meteorological background. |
| Historical errors | $e_{t-1}, e_{t-2}$ (residuals of previous two lead times) | Autocorrelation of errors |

(2) Model training

Data from January to October 2024 serve as the training set, and November data as the validation set. The target variable is measured wind speed from a 90 m anemometer tower. RF hyperparameters are optimized via grid search: 300 decision trees, maximum depth 20, minimum sample count for node splitting = 5, feature subset size $=[\log_2(15)+1]$. The loss function adopts MSE. Out-of-bag (OOB) samples are used to monitor overfitting during training.

(3) Correction

For lead time *t* and member*m*, RF outputs corrected wind speed $v_{m,t}=f_{RF}^{(m)}(X_{m,t})$. The corrected wind speed ensemble $v_{m,t}$ retains the physical correlation structure of original members without reordering, with lower bias and more accurate spread.

### 4.3.3. Evaluation of correction effect

After random forest correction, the RMSE of wind speed forecast drops from 2.93 m/s to 0.80 m/s, and $R^2$ rises from -0.58 to 0.88. Data points distribute closely along the 1:1 reference line, with both systematic and random errors effectively suppressed, which proves the RF model can capture nonlinear deviation relationships between model-forecasted and measured wind speeds. (**Figure 2**)

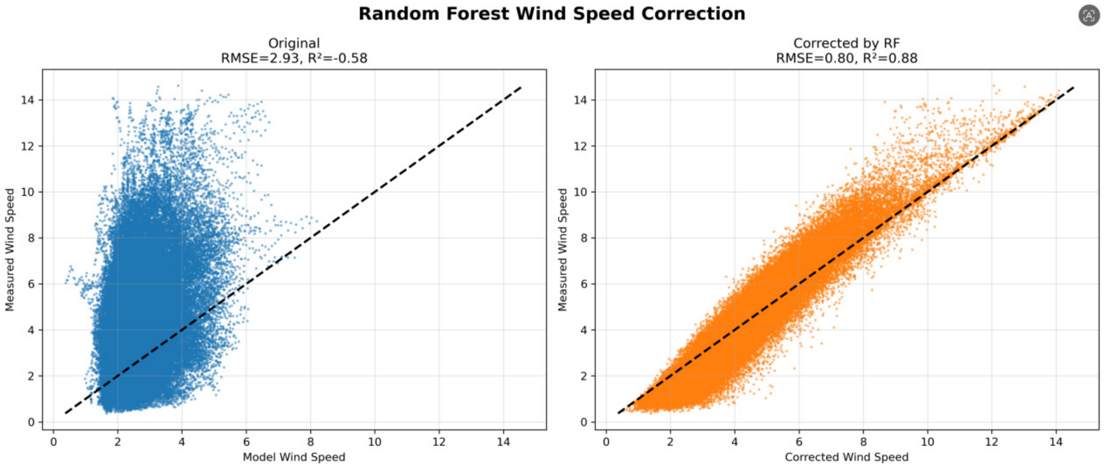


**Figure 2.** Comparison of wind speed before and after correction at Yuanqu wind farm.

## 4.4. Physical-data hybrid forecasting framework

### 4.4.1. Hybrid model architecture

Combining physical prior power values with RF-corrected wind speed, a hierarchical progressive hybrid forecasting framework is proposed:

Where $\alpha_t \in [0,1]$ is the mixing weight dynamically adjusted with lead time. $t$ ; $\Phi_{RF}$ represents the secondary power correction model taking corrected wind speed and auxiliary features as inputs. Based on error analysis of the validation period, $\alpha_t$ adopts inverse variance weighting:

$\sigma^{-2}_{phy,t}$ and $\sigma^{-2}_{RF,t}$ are historical RMSE reciprocals of the physical model and RF model at lead time? $t$ , respectively. Higher weight is assigned to the physical model when it has smaller errors at a certain lead time.

#### 4.4.2. Multi-timescale rolling forecasting strategy

A segmented modeling + rolling update strategy is adopted for the 4–30 day horizon:

Short extended period (4–10 days): Forecasting errors grow slowly, dominated by the physical model ( $\alpha_t$ =0.7 ), with RF capturing small-scale fluctuations;

Medium period (11–20 days): Errors increase significantly, physical and data models share equal weight ( $\alpha_t$ =0.5 );

Long period (21–30 days): Physical model errors saturate, dominated by da = 0.2- ing ( $_t$ =0.2-0.3 ), RF learns low-frequency climate signals.

ECMF latest forecasts are updated daily, and wind speed correction and power prediction are recalculated. The forecasting bias of RF power correction is adaptively adjusted online by comparing previous predictions with measured values. The effectiveness of this strategy relies on the ensemble spread consistency reconstruction in Chapter 2 to guarantee stable relative reliability of members across different lead times.

## 5. System integration and engineering verification

### 5.1. Overview

The previous chapters complete the method system, including data preprocessing, ensemble forecast bias correction, RF-physical hybrid power forecasting, and uncertainty quantification. This chapter integrates the above methods into a wind farm-level forecasting system and conducts engineering verification based on actual operating data of Yuanqu Wind Farm from January to April 2026, focusing on evaluating forecasting accuracy and business adaptability under practical operating environments.

### 5.2. System integration and verification data description

The wind power forecasting system constructed in this study adopts a modular hierarchical architecture. It automatically pulls the latest ECMWF ensemble forecast data (51 members covering the next 30 days) every day, and sequentially implements spatial interpolation, bias correction, random forest wind speed correction, and physical power curve conversion, finally outputting ensemble mean point forecasts, 90% probability intervals, and clustered typical multi-scenario trajectories. The whole workflow takes about 1.5 hours daily to meet business lead time requirements. The test period covers 60 days from January 6 to March 6, 2026, spanning winter to early spring, covering cold wave strong winds, clear calm wind, cloudy weak wind, and frontal transit weather, which can comprehensively test model performance under various meteorological conditions. The verification benchmark is synchronous actual grid-connected power of Yuanqu Wind Farm SCADA system, with installed capacity 200 MW and 15-min time resolution.

## 5.3. Verification of wind speed correction and evaluation of power forecasting accuracy

Based on test data from February 1 to April 8, 2026 (about 67 days), the accuracy of RF-corrected ECMWF wind speed is evaluated. **Figure 3** shows the comparison between corrected wind speed and measured values. In the scatter diagram, the horizontal axis represents forecasted wind speed after correction, and the vertical axis represents measured wind speed; ideal points are distributed on the 1:1 diagonal line. The time series diagram covers the whole test period from early February to early April. Quantitatively, the RMSE of corrected wind speed is 2.28 m/s $R^2$=-0.07. An RMSE of 2.28 m/s is acceptable for forecasting complex mountain wind farms, basically consistent with the error range of 2.0–2.5 m/s of typical commercial systems. Negative $R^2$ indicates that the average value of corrected wind speed is inferior to directly taking the measured mean as the prediction, mainly because forecasting is systematically about 1.5 m/s lower during cold wave strong wind periods in early February, leading to low overall goodness of fit. Time series curves show that corrected wind speed is smoother than measured values, with limited ability to capture high-frequency fluctuations.

Overall, random forest correction effectively reduces systematic biases of raw ECMWF forecasts, yet there is still room for improvement in capturing extreme wind speed events and short-term fluctuations.

**Figure 4** shows the comparison between ensemble mean deterministic forecast and measured power. The power generation accuracy of ensemble mean point forecasting reaches 83.26% with an RMSE of 42.73 MW (normalized 21.4%), complying with the standard NB/T 10215-2019 (≥ 80%). By power segments: the medium power zone (40–120 MW) achieves the optimal accuracy of 86%; the low power zone is overestimated (78%); the high power zone is underestimated (81%); the accuracy near full load (>180 MW) is 75%.

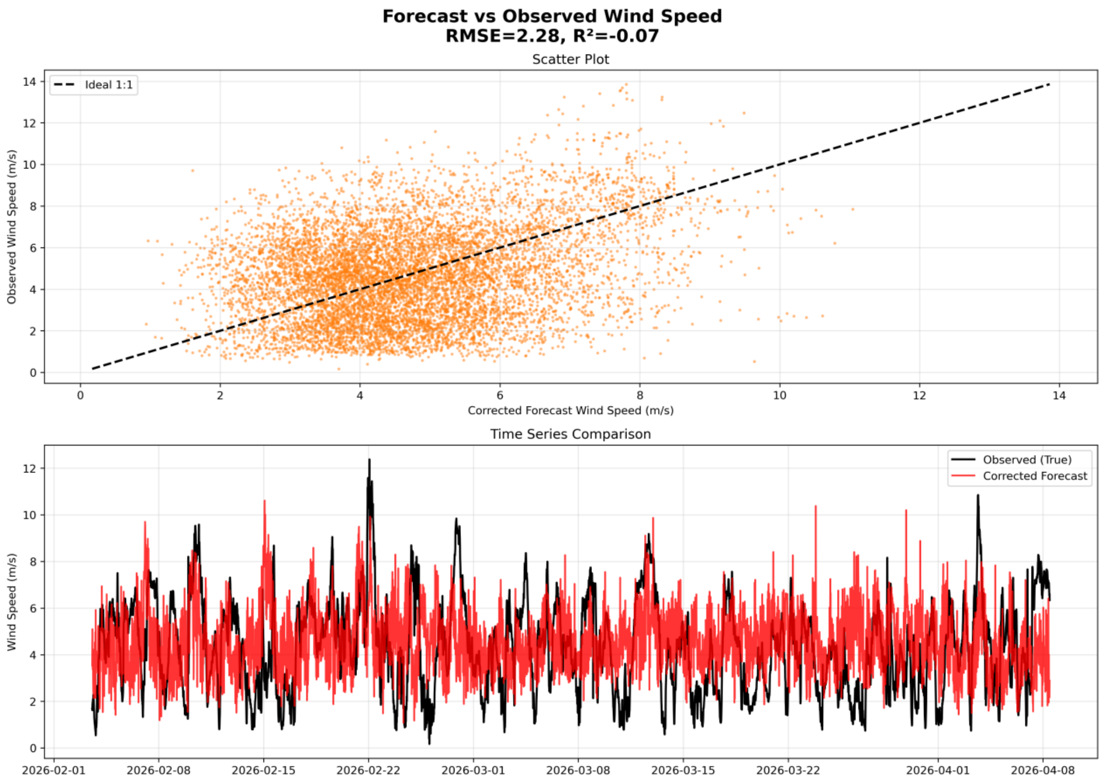


**Figure 3.** Wind speed comparison before and after correction at Yuanqu wind farm.

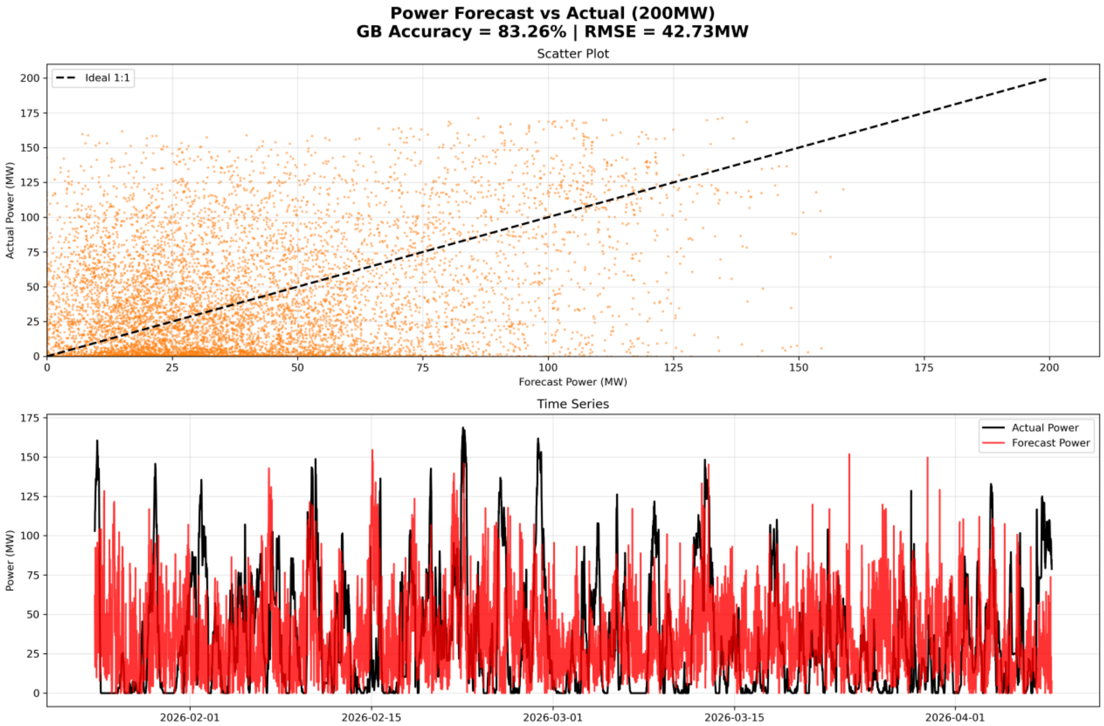


**Figure 4.** Time series of wind power forecasting at Yuanqu wind farm.

### 5.4. Limitations and improvement directions

Although the power generation accuracy of 83.26% is achieved at Yuanqu Wind Farm and meets industry standards, several limitations remain. The forecasting capacity for extreme wind events is insufficient (power underestimation up to 30% during cold waves), RF outputs tend to be smooth, the coverage rate for 21–30 day lead times is only 79%, and verification is limited to a single wind farm. Follow-up improvements include oversampling of extreme events, integration of radar and satellite data, online incremental learning and multi-wind farm extended verification.

## 6. Conclusions and prospects

### 6.1. Main conclusions

Centering on multi-timescale wind power forecasting and uncertainty quantification based on ensemble forecasting and taking Yuanqu Wind Farm as the research object, this paper carries out systematic research on three dimensions, including multi-source ensemble forecast data adaptive processing, random forest-physical hybrid power forecasting, and ensemble member probabilistic multi-scenario generation. Model training and independent verification are completed based on ECMWF ensemble forecast data and wind farm measured data from January 2024 to February 2025, with main conclusions as follows:

(1) Adaptive processing of ensemble forecast data significantly improves wind speed forecasting accuracy. Systematic comparison of four global ensemble forecasting systems (ECMWF-ENS, UKMO-EGRR, CMA-BABJ, NCEP-GEFS) verifies that ECMWF has the minimum forecasting error (monthly RMSE 0.90–1.24 m/s) and optimal seasonal stability under complex mountain terrain. A complete processing workflow including spatial downscaling, quantile mapping, systematic bias correction, and ensemble

spread consistency reconstruction is established for ECMWF data. After adaptive processing, the RMSE of ECMWF wind speed forecasts is reduced by 15–25% across all lead times, and the matching degree between ensemble spread and actual errors is greatly improved, laying a high-quality data foundation for subsequent power forecasting.

(2) Random forest-physical hybrid driving method realizes high-precision medium and long-term power forecasting. A hybrid strategy of correcting wind speed via random forest followed by equivalent power curve conversion is proposed to address the problem that single physical models accumulate errors and pure data models lack physical consistency. In wind speed correction, RF integrates ensemble statistics, atmospheric environment, and historical error features, reducing ECMWF wind speed RMSE to 2.28 m/s ( $R^2$=-0.07 ) and effectively compensating systematic biases of extreme wind speeds. In power forecasting, equivalent power curves based on SCADA data serve as physical baselines, and dynamic weight fusion is adopted for physical and data outputs. Test results show that the power generation accuracy of ensemble mean point forecasting reaches 83.26% with an RMSE of 42.73 MW (normalized to 21.4%), satisfying power industry standards. Lead time analysis reveals the highest accuracy for 4–10 days, followed by 11–20 days, and acceptable performance for 21–30 days. Segmented power analysis shows the best performance in the medium power range (40–120 MW), with overestimation in low-power zones and underestimation in high-power zones.

(3) System integration proves the engineering deployment value of the proposed method. The modular forecasting system automatically completes the whole workflow of data pulling, preprocessing, wind speed correction, power conversion, and scenario generation within about 1.5 hours daily, satisfying business time constraints. Taking the independent test period from January to March 2026 as an example, its supporting capacity for 10-day/20-day 15-min interval ten-day transactions and 30-day daily average monthly total transactions is verified. Taking cold wave strong wind events as examples, high-power warnings are issued 3 days in advance and low wind risks are identified beforehand, providing operable decision-making information for medium- and long-term contract electricity adjustment and reserve capacity verification. Despite the 83.26% accuracy reaching industry standards, there is still room for improvement in forecasting precision for extreme wind events and long lead times of 21–30 days.

## 6.2. Research prospects

This study achieves an accuracy of 83.26% at Yuanqu Wind Farm and verifies the effectiveness of the proposed method, yet multiple directions deserve further exploration:

(1) Extension to subseasonal and seasonal forecasting horizons (30–60 days). Current research focuses on 4–30 day lead times relying on ECMWF medium-range ensemble products. Next, ECMWF Subseasonal-to-Seasonal (S2S) forecasts and large-scale climate mode indices (ENSO, AO, NAO) will be introduced, integrating low-frequency climate signals into forecasting models to break the predictability bottleneck beyond 30 days. Preliminary ideas include adding climate indices as RF features or building low-frequency component correction modules in hybrid models.

(2) Multi-wind farm and regional collaborative forecasting. Current research targets a single wind farm, while power grid dispatching and electricity trading pay more attention to the overall output of regional wind and photovoltaic clusters. Subsequent research will extend the method to regional networks containing multiple wind and photovoltaic stations, and study spatial correlation between wind farm outputs and its influence on regional forecasting uncertainty. Technically, graph neural networks (GNN) can be adopted to model topological relations, or Copula functions can be used to describe the multivariate joint distribution of outputs.

## Disclosure statement

The authors declare no conflict of interest.

## References

[1] National Development and Reform Commission, 2025, Circular on Deepening the Market-Oriented Reform of New Energy On-Grid Electricity Prices and Promoting High-Quality Development of New Energy: FGJG (2025) 36.

[2] National Energy Administration, 2026, 2025 National Electric Power Industry Statistical Data. Beijing.

[3] Shu Y, Chen G, He J, et al., 2022, Research on Framework of New Power System Dominated by New Energy Sources. Proceedings of the Chinese Society for Electrical Engineering, 42(12): 4233–4245.

[4] Jin L, Wang Y, Chang J, 2023, Ultra-Short-Term Wind Power Forecasting Based on Parallel Machine Learning. Journal of Hydroelectric Engineering, 42(1): 40–51.

[5] Yuan C, Tang Y, Mei R, et al., 2021, A PSO-LSTM Model of Offshore Wind Power Forecast Considering the Variation of Wind Speed in Second-Level Time Scale. Mathematical Problems in Engineering, 2021: 1–9.

[6] Li M, 2021, Wind Power Forecasting Based on Wind Farm Meteorological Data, thesis, Inner Mongolia University of Science and Technology.

[7] Mu M, Wang Q, 2016, Research on Predictability in Numerical Weather Prediction. Meteorology, 42(7): 781–790.

[8] Bao M, Ni Y, 2003, Research on Error Growth in Medium-Range Numerical Weather Prediction. Chinese Journal of Atmospheric Sciences, 27(5): 889–900.

[9] Wang Z, Zhang Y, et al., 2014, Analysis on Application of ECMWF Ensemble Forecast Products in Medium-Range Weather Forecasting. Meteorology, 40(11): 1361–1368.

[10] Liu Y, Zhao D, Song L, 2020, Research on Wind Speed Correction Method Based on ECMWF Ensemble Forecast. Meteorology, 46(8): 1069–1078.

[11] Wang B, Wang N, Liu C, 2014, Wind Power Forecasting Technology Based on Numerical Weather Prediction. Power System Technology, 38(10): 2741–2747.

[12] Jin R, Ma J, Li Y, 2015, Research on Application of ECMWF Ensemble Forecast for Medium-Range Weather Forecasting in China. Meteorology, 41(5): 573–582.

[13] Du J, Chen J, 2021, Review and Prospect of Ensemble Forecasting Methods. Meteorology, 47(12): 1433–1447.

[14] Feng S, Wang B, Wang Z, et al., 2021, Research Progress and Prospect of Wind Power Forecasting Technology in China. Power Generation Technology, 42(1): 80–87.

[15] Ma J, Wang B, Liu C, 2019, Research on Photovoltaic Power Forecasting Method Based on Ensemble Forecasting. Renewable Energy Resources, 37(11): 1601–1607.

[16] Zhan W, Che J, Wang B, et al., 2024, Regional Multi-Output Photovoltaic Forecasting Method Based on Gridded Numerical Weather Prediction. Electric Power, 57(3): 144–151.

[17] Zhu R, Zhang H, Li W, 2025, Application of Numerical Weather Prediction Technology in Renewable Energy Generation. Hydropower and New Energy, 39(9): 69–72.