

KNOWLEDGE-ENHANCED XGBOOST FOR WIND POWER FORECASTING UNDER EXTREME METEOROLOGICAL EVENTS

Jin Ma^{}, Ming Xu, Wei Du, Zhongzhi Qiao

Received on March 12, 2026

Presented by K. Stoykova, Corresponding Member of BAS, on June 30, 2026

Abstract

To address deteriorated generalization of conventional data-driven wind power forecasting under extreme meteorological conditions, this study develops a knowledge-enhanced XGBoost prediction framework. Physical thresholds are used to divide meteorological and turbine control variables into state labels for wind speed, temperature, yaw angle, and pitch angle. The labels are compressed into two-dimensional numerical features and combined with sliding window temporal features for model training. Experiments based on the SDWPF dataset verify that embedded physical prior information mitigates out-of-distribution prediction drift. Compared with vanilla XGBoost, the proposed method reduces 1-hour MAE by 23.77% with statistical significance and achieves evident accuracy improvement across diverse extreme weather subsets.

Key words: extreme weather events, event labels, knowledge injection, wind power forecasting

1. Introduction. Under the background of the dual carbon goals and energy transition, wind power is rapidly integrating into the main grid structure. However, its output is significantly affected by meteorological conditions and exhibits typical characteristics such as randomness, intermittency, and nonlinearity [1, 2]. Under the influence of extreme weather, the operating state of wind turbines often

This research was funded by State Grid Shanxi Electric Power Co., Ltd. Technology Project OF FUNDER grant number 5205D024000B.

<https://doi.org/10.7546/CRABS.2026.07.05>

undergoes qualitative changes. These critical meteorological conditions cause the wind turbine to trigger cut-out protection, power derating, or a significant change in aerodynamic performance. These conditions specifically cover low-temperature freezing, strong wind cut-out, high-temperature heat waves, and ramp events [3]. Although such events account for a limited proportion, they significantly elevate the non-stationarity of power sequences, making models trained on historical statistical rules more prone to lags and deviations [4]. Therefore, accounting for extreme weather events to improve the accuracy and robustness of wind power forecasting is of great significance for building a resilient grid [5].

Existing wind power forecasting methods mainly include physical models, statistical models, and data-driven models [6–8]. However, when facing extreme weather events, such methods encounter two core technical bottlenecks. The first is the long-tail effect of data distribution. Objective functions based on global error minimization are often dominated by normal samples, leading the model to sacrifice the accuracy of sparse extreme samples. The second is the multi-modal switching of input-output mapping relationships. Under special working conditions, the operating mechanism of the turbine is prone to qualitative changes, causing the power response characteristics to be completely different from normal conditions [9].

Extreme Gradient Boosting (XGBoost) uses gradient boosting trees as its core and is widely used in short-term wind power forecasting [10]. However, traditional XGBoost still has limitations when dealing with extreme weather problems. Its essence is a greedy algorithm based on feature value splitting, making it difficult to automatically capture the semantic mutations of meteorological variables under specific physical thresholds. Therefore, solving the aforementioned generalization challenges hinges on explicitly sensing extreme events and effectively integrating physical-level event knowledge into the data-level regression learning process.

To this end, this study proposes an XGBoost-based model enhanced with extreme-event knowledge features. The core academic contributions and innovations of this work are as follows:

1. Targeting the multi-modal switching problem of wind turbine operation caused by extreme weather, an explicit knowledge injection mechanism based on physical thresholds is proposed.
2. To address the underfitting problem of long-tail distributions caused by the scarcity of extreme event samples, a hybrid feature space combining sliding window and event encoding is designed.
3. Multi-horizon and sub-scenario experiments on the real SDWPF [11] dataset quantitatively proved that introducing prior physical rules can significantly suppress the performance degradation of pure data models in out-of-distribution scenarios.

2. Proposed model. To improve the stability and generalization performance of wind power prediction under complex meteorology, the proposed method employs the standard XGBoost as the backbone of the regression model. Through event segmentation, label encoding, knowledge feature injection, and an integrated prediction pipeline, the physical state information is incorporated into the model input.

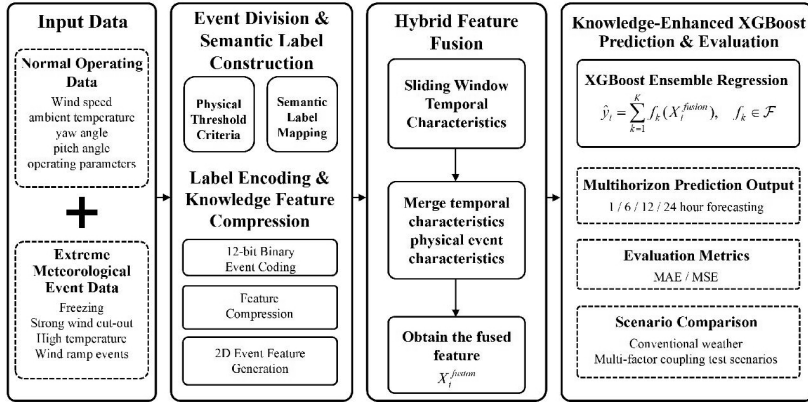


Fig. 1. The overall framework of the model

2.1. Overall framework of the model. As shown in Fig. 1, the proposed model follows a sequential workflow from multi-source input data, event-based feature construction, hybrid feature fusion, to knowledge-enhanced XGBoost prediction and evaluation.

2.2. Event interval division and semantic label construction. To address the limitations of solely relying on continuous numerical features, an event interval division and label construction mechanism is introduced. Table 1 details the semantic interval divisions and label mappings.

2.3. Label feature encoding and knowledge injection. To characterize both normal fluctuations and abrupt changes in a unified framework, event label features are introduced in knowledge-enhanced XGBoost, encoding state information into structured semantic quantities. A 12-bit binary event code is defined for each moment. To avoid sparsity and dimensional redundancy, this is compressed into a 2-dimensional continuous semantic feature:

$$(1) \quad E_{weather} = \sum_{i=1}^6 b_i \cdot 2^{6-i}$$

$$(2) \quad E_{control} = \sum_{i=7}^{12} b_i \cdot 2^{12-i}$$

$$(3) \quad E_{event} = [E_{weather}, E_{control}]$$

T a b l e 1

Semantic interval division of main physical features and label mapping

Feature Type	Feature Variable	Interval Condition	Semantic Label	Encode
Meteorological	Wind speed	$v < 2$	Low wind	b_1
		$2 \leq v < 15$	Normal wind	b_2
		$v \geq 15$	Strong wind	b_3
	Ambient temperature	$T_a < -5$	Freezing	b_4
		$-5 \leq T_a \leq 35$	Moderate	b_5
		$T_a > 35$	Hot temperature	b_6
Control	Yaw angle	$ \theta < 10^\circ$	Aligned	b_7
		$10^\circ \leq \theta \leq 25^\circ$	Slight misalignment	b_8
		$ \theta > 25^\circ$	Yaw misalignment	b_9
	Average pitch angle	$\bar{\beta} < 5^\circ$	Full-load	b_{10}
		$5^\circ \leq \bar{\beta} \leq 10^\circ$	Pitch control	b_{11}
		$\bar{\beta} \geq 10^\circ$	Power-limited	b_{12}

By appending E_{event} to the end of the sample features, event information is embedded into every training sample and participates in regression learning alongside numerical features.

2.4. Wind power prediction. Based on the fused features, the power prediction model uses XGBoost for nonlinear regression mapping. Assuming the model contains K trees, the overall output is:

$$(4) \quad \hat{y}_i = \sum_{k=1}^K f_k(X_i^{fusion}), \quad f_k \in \mathcal{F}$$

where $\mathcal{F}$ is the function space of all regression trees. The method minimizes the regularized objective function:

$$(5) \quad Obj = \sum_{i=1}^N L(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k),$$

where $\Omega(f_k)$ is the model complexity penalty term to suppress overfitting.

3. Experiments and analysis. 3.1. Data source and overall features. Experimental data derives from the public SDWPF dataset, containing 2020–2021 SCADA operating data and NWP variables for 134 wind turbines in a northern China wind farm, with a 30-minute time resolution. Ten key variables were selected based on wind energy conversion mechanisms, including Wind speed (W_{spd}), Wind direction angle (W_{dir}), Ambient temperature (E_{tmp}), Nacelle temperature (I_{tmp}), Nacelle direction (N_{dir}), Pitch angle ($P_{ab1}, P_{ab2}, P_{ab3}$), Reactive power (P_{rtv}) and Active power (P_{atv}).

T a b l e 2

Quantitative results of a full sample comparative experiment

Metric	Model	1-h	6-h	12-h	1-d
MAE	SVM	94.33±7.40***	187.01±16.25***	246.29±21.52***	303.28±27.46***
	LR	167.32±32.46***	201.68±27.80***	252.66±30.24***	305.87±36.50**
	KNN	113.75±11.32***	214.07±20.31***	278.30±31.64***	333.64±32.82***
	LSTM	87.95±9.49***	213.88±27.50***	282.57±35.90***	345.07±34.81***
	XGBoost	83.75±8.50***	172.19±16.74**	233.68±22.53*	295.94±29.73**
	Ours	63.84±5.53	160.81±16.32	225.90±24.88	285.86±32.67
MSE	SVM	23.50±3.68***	74.19±11.17***	119.66±18.87***	173.17±27.24***
	LR	76.41±84.71*	82.80±39.12**	105.72±27.87*	141.89±31.25*
	KNN	40.86±8.19***	106.30±18.55***	159.81±31.55***	208.82±40.06***
	LSTM	24.70±6.04***	104.52±26.17***	157.49±37.70***	220.52±43.61***
	XGBoost	18.72±3.71***	55.00±10.13**	90.20±16.91*	132.34±36.99*
	Ours	12.75±2.60	52.47±8.70	88.48±16.63	132.03±27.06

Note: Mean ± standard deviation. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, compared with Ours.

3.2. Experimental setup and parameter configuration. To ensure reproducible and consistent comparisons across models, we randomly select 10 wind turbines from the original dataset for performance evaluation, rather than a single turbine, to improve the statistical significance of experimental results. The historical time step $P = 16$ (corresponding to 8 hours), and the prediction steps $T = 2, 12, 24, 48$, corresponding to the 1-hour, 6-hour, 12-hour, and 1-day prediction tasks, respectively.

We configure adaptive hyperparameters for the XGBoost regressor according to different prediction horizons. When the prediction step $T \leq 12$, the learning rate is set to 0.015, the maximum tree depth is 7, the subsample ratio and column sample ratio are both set to 0.85, the minimum child weight is 3, the L2 regularization coefficient is 1.0, the minimum loss reduction is 0.1, and the early stopping round is 100. For the condition of $12 < T \leq 24$, we adjust the maximum tree depth to 5, the column sample ratio to 0.70, the minimum child weight to 5 and the L2 regularization coefficient to 5.0 while keeping the learning rate unchanged. For long-term prediction with $T > 24$, we further reduce the learning rate to 0.010, the maximum tree depth to 3, the subsample ratio to 0.70 and the column sample ratio to 0.50, and increase the minimum child weight to 15, the L2 regularization coefficient to 15.0, the minimum loss reduction to 1.0 and the early stopping round to 200.

Furthermore, this work divides the extreme scenarios into four basic events: high temperature, low temperature, strong wind and weak wind, and further constructs a coupled event with simultaneous triggering of wind speed and temperature to characterize the superposition effect of multiple factors.

3.3. Experimental metrics and baseline methods. To comprehensively evaluate the model performance, two metrics, Mean Absolute Error (MAE) and Mean Squared Error (MSE), are adopted:

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

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

3.3.1. Full-sample prediction. Table 2 reveals that all baseline models produce statistically significant larger prediction errors. Compared with standard XGBoost, the proposed model reduces one-hour MAE by 23.77% and one-day MSE by 0.24%.

3.3.2. Prediction under extreme meteorological scenarios. As demonstrated by the quantitative results in Table 3, our model exhibits superior forecasting robustness and accuracy across various extreme meteorological conditions, including strong winds, abnormal temperatures, and multiple factor coupled scenarios,

T a b l e 3

Quantitative results of the comparative experiment on extreme weather samples

Type	Time steps	SVM		LR		KNN		LSTM		XGBoost		Ours	
		MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
High temperature	1-h	77.21	20.84	162.96	50.52	136.60	40.52	146.89	43.77	77.07	9.70	71.37	9.08
	6-h	166.29	80.20	335.45	182.15	221.86	92.03	265.31	117.35	157.45	40.91	150.22	37.93
	12-h	197.21	101.78	377.39	211.84	250.13	109.86	263.79	119.70	215.81	69.88	195.16	60.38
	1-d	276.45	109.42	386.00	214.40	246.85	106.53	259.87	110.56	275.81	108.65	256.61	100.00
Low temperature	1-h	527.05	381.34	216.61	77.33	287.97	98.44	160.99	54.13	605.40	560.88	594.89	559.71
	6-h	506.68	351.77	1079.02	1522.38	299.76	108.82	235.47	76.65	567.77	402.46	564.83	401.08
	12-h	593.44	398.23	565.73	370.86	682.24	609.70	556.29	358.04	630.01	464.03	615.31	454.58
	1-d	681.66	409.67	779.57	667.67	467.80	269.69	297.62	99.24	632.90	462.51	637.53	489.98
Strong wind	1-h	129.96	78.28	335.45	182.15	103.67	37.74	96.48	26.82	115.82	63.98	111.69	64.13
	6-h	238.19	151.95	406.71	271.61	171.46	95.10	220.39	99.37	172.83	63.44	158.26	53.03
	12-h	361.37	206.34	566.60	361.30	281.46	117.40	369.87	206.38	260.22	107.45	262.34	109.16
	1-d	869.29	910.28	926.06	888.29	606.18	505.04	386.30	282.95	644.22	471.84	638.43	465.31
Weak wind	1-h	89.37	23.92	534.76	207.45	72.64	16.26	67.22	9.23	73.94	8.07	38.79	4.33
	6-h	194.75	82.34	1837.28	2291.65	180.52	72.62	265.17	140.60	203.98	61.47	197.73	60.14
	12-h	274.49	118.84	2265.60	2870.03	236.90	112.10	266.69	126.05	309.34	124.46	269.86	101.51
	1-d	332.69	180.22	1071.25	5745.85	368.02	141.39	768.05	705.94	366.41	156.62	320.00	126.14
Coupling	1-h	51.53	11.83	123.26	40.92	118.11	29.84	117.57	25.27	78.28	8.59	52.24	5.44
	6-h	273.44	74.52	368.24	180.49	264.67	114.67	328.17	157.23	214.80	57.11	204.56	52.30
	12-h	276.97	111.12	448.09	239.32	280.01	126.55	286.95	126.44	317.04	117.80	273.25	92.33
	1-d	417.93	234.06	489.53	263.02	387.85	137.93	349.96	151.05	378.15	157.95	343.30	134.59

because its knowledge injection mechanism explicitly incorporates physical operational boundaries to effectively overcome the out of distribution generalization failures typical of traditional pure numerical fitting models.

4. Conclusions. To improve wind power forecasting accuracy under complex meteorological conditions, this work proposes an XGBoost-based model enhanced with extreme-event knowledge features. By preserving time dependencies through sliding windows and explicitly mapping critical physical states into event labels, our model effectively bridges data-driven capabilities with physical operational logic. Experiments on the SDWPF dataset verified its superiority, notably showing significant error reductions and enhanced robustness across normal operation, extreme weather events, and complex multi-factor coupled scenarios. Future work will refine event libraries and introduce adaptive weighting mechanisms for varied disturbance magnitudes.

REFERENCES

- [1] LI M., M. YANG, Y. YU et al. (2024) Short-term wind power forecast based on continuous conditional random field, *IEEE Trans. Power Syst.*, **39**(1), 2185–2197.
- [2] LIU Z., F. DING, J. LU et al. (2022) Comparison of BLSTM-attention and BLSTM-transformer models for wind speed prediction, *C. R. Acad. Bulg. Sci.*, **75**(1), 80–89.
- [3] CHANG Y., H. YANG, Y. CHEN et al. (2024) A hybrid model for long-term wind power forecasting utilizing NWP subsequence correction and multi-scale deep learning regression methods, *IEEE Trans. Sustain. Energy*, **15**(1), 263–275.
- [4] FARAHBOD S., T. NIKNAM, M. MOHAMMADI et al. (2022) Probabilistic and deterministic wind speed prediction: ensemble statistical deep regression network, *IEEE Access*, **10**, 47063–47075.
- [5] BABU K. V. S. M., D. DWIVEDI, P. CHAKRABORTY et al. (2024) A resilient power distribution system using P2P energy sharing, *IEEE Trans. Ind. Appl.*, **60**(6), 8228–8238.
- [6] PRAKASH A., S. Y. LEE, X. LIU et al. (2024) A Bayesian hierarchical model to understand the effect of terrain on wind turbine power curves, *IEEE Trans. Sustain. Energy*, **15**(2), 1127–1137.
- [7] ZHANG G., Z. LI, K. ZHANG et al. (2019) Multi-objective interval prediction of wind power based on conditional copula function, *J. Mod. Power Syst. Clean Energy*, **7**(4), 802–812.
- [8] TIAN J., G. ZENG, J. ZHAO et al. (2024) A data-driven modeling method of virtual synchronous generator based on LSTM neural network, *IEEE Trans. Ind. Inform.*, **20**(4), 5428–5439.
- [9] YANG Z., X. PENG, J. SONG et al. (2024) Short-term wind power prediction based on multi-parameters similarity wind process matching and weighed-voting-based deep learning model selection, *IEEE Trans. Power Syst.*, **39**(1), 2129–2142.
- [10] KEERTHISINGHE C., A. R. SILVA, P. TARDAGUILA et al. (2023) Improved short-term wind power forecasts: low-latency feedback error correction using ramp pre-

diction and data from nearby farms, IEEE Access, **11**, 128697–128705.

- [11] ZHOU J., X. LU, Y. XIAO et al. (2024) SDWPF: A dataset for spatial dynamic wind power forecasting over a large turbine array, Sci. Data, **11**, 649.

*State Grid Shanxi Electric Power Co., Ltd. Changzhi Power Supply Branch,
63 East Taihang Street, Luzhou District, Changzhi City, Shanxi Province, China*
e-mails: mjcزدgs@163.com, <https://orcid.org/0009-0006-9346-4032>
xmcزدgs@163.com, dwcزدgs@163.com, qzzcزدgs@163.com