Prediction Of Wind Power Using An Optimized Ensemble Machine Learning Model

Arangarajan Vinayagam[1*], Kavitha M V[2], Deepa A[3], Senthil Kumar H[4], Arivoli Sundaramurthy[5], and Saravanan K[6]

[1] Department of Electrical and Electronics Engineering, New Horizon College of Engineering, Bangalore, India
[2] Department of Electronics and Communication Engineering, Cambridge Institute of Technology, Bangalore, India
[3] Department of Electronics and Communication Engineering, Gopalan College of Engineering and Management, Bangalore, India
[4] Department of Computer Science and Engineering, Presidency University, Bangalore, India
[5] Department of Electrical and Electronics Engineering, PSG Institute of Technology and Applied Research, Coimbatore, India
[6] Department of Electrical and Electronics Engineering, SRM Institute of Science and Technology, Kattankulathur-603203, Chengalpattu, Tamil Nadu, India
[*] Corresponding author. E-mail: arajanin@gmail.com

Received: May 07, 2026; Accepted: Jun. 02, 2026

Wind power forecasting is very crucial to ensure stable renewable energy (RE) integration and the stability of power systems. In this research work, several machine learning (ML) models, such as linear regression (LR), random forest (RF), gradient boosting (GB), extreme GB (XGBoost), and grid search cross-validation (GS-CV) optimized XFGBoost are analyzed in terms of their performance for wind power forecasting using a real-life large dataset having weather and temporal features. Performance of these models is analyzed using metrics such as mean absolute error (MAE), root mean square error (RMSE), and correlation factor (R^2). It is found that XGBoost with GS-CV approach consistently performs better than other models with the minimum prediction error and maximum R^2 value. The results show that hyperparameter-optimized XGBoost can greatly enhance forecasting accuracy and generalization, providing a practical data-driven solution for wind power forecasting and decision-making.

Keywords: Renewable Energy, Machine Learning, Gradient Boosting, Extreme Gradient Boosting, Random Forest, Wind Power Prediction

http://dx.doi.org/10.6180/jase.202611_34.050

1. Introduction

The nonlinear and volatile nature of wind energy imposes significant challenges for modeling and prediction [1]. Conventional statistical methods, like autoregressive moving average (ARIMA) models and their variants, are commonly used for wind energy prediction due to their simplicity, interpretability, and ease of implementation [2]. While these methods are effective for short-term analysis, their accuracy decreases for longer-term forecasts, particularly when the process is highly nonlinear and variable [3]. Basic machine learning (ML) models, including Support Vector Regressor (SVR) [4], decision tree (DT) Regressor, k-nearest neighbors (KNN) [5], Artificial Neural Network (ANN), Multi-layer Perceptron (MLP) [6], and Radial Basis Function (RBF) [7] have been introduced to address the challenges existing with traditional statistical models. However, the basic ML models have drawbacks of their own. These include a loss of interpretability, a tendency to overfit, and a need for proper feature selection and scaling, particularly when dealing with large, turbulent data sets [7]. On the other hand, many researchers have employed ensemble tech-

Table 1. A sample of Wind Power Dataset

Date/Time	LV Active Power (kW)	Wind Speed (m/sec)	Theoretical power curve (kWh)	Wind direction (°)
01-01-2018 00:00:00	380.047.	5.311	416.328	259.994
01-01-2018 00:10:00	453.769.	5.672	519.917	268.641
01-01-2018 00:20:00	306.376	5.216	390.900	272.564

niques in recent research to enhance the performance, robustness, and practical value of multiple weak learners [8]. Random Forest (RF) [9], Stacking ensemble [10], and Gradient-Boosting methods, including CatBoost, XGBoost, and LightGBM, have shown superior performance in short- and medium-term wind power prediction [11]. Among ensemble methods, XGBoost has emerged as a preferred technique for predicting wind power, thanks to its regularized gradient-boosting, rapid parallel tree building, and impressive performance on meteorological data [12].

Moreover, for forecasting wind energy, various optimizations have been proposed to increase the efficiency of XGBoost, such as Bayesian optimization [13], variational ant colony optimization (ACO) [14], genetic algorithm (GA) optimization [15] and the grid search with cross-validation method (GS-CV) [16]. Out of all the above optimizations, GS-CV proves to be more reliable with high prediction accuracies and reportable parameters, which makes it more desirable in the case of XGBoost wind power forecasting applications because of its simplicity and systematic nature. However, the GS-CV method is used in only a few studies, as most studies use either metaheuristic or Bayesian methods with opaque tuning. Hence, this work adopts GS-CV-optimized XGBoost and explores a defined hyperparameter space (number of trees, maximum depth, learning rate, subsampling, and column-sample ratios) in a deterministic manner. Using K-fold cross-validation minimizes overfitting and establishes a clear optimization protocol. It outperforms LR, GB, RF, and standard XGBoost models, with lower error metrics (MAE and RMSE) and a higher R^2. The findings show that systematic hyperparameter optimization significantly improves the accuracy of XGBoost's short-term wind power estimates. The main contributions of this study are as follows: (i) Forecasting wind power generation is done with utilization of LR and ensemble predictive models (GB, XGBoost, and CS-CV optimized XGBoost); (ii) An optimized XGBoost predictive model is developed using the GS-CV optimization strategy; and (iii) The predictive models are trained, and the predictive metrics (MAE, RMSE, and R^2) are evaluated to confirm their performance.

2. Materials and methods

For predicting the wind power generation, the ML algorithms such as LR, GB, XGBoost, RF, and GS-CV-optimized XGBoost are utilized in the present study. The models are trained using real-time wind power data obtained from Kaggle [17]. The dataset comprises SCADA data recorded at 10-minute intervals from 1 January 2018 to 31 December 2018, consisting of four attributes: wind speed, theoretical power curve, wind direction, and wind power. For example, the top three rows of the dataset variables are shown in Table 1.

The ML algorithms were designed and evaluated using the Python 3.10 version in the Google Colaboratory platform, whereas simple data processing and basic data analysis were conducted on the personal computer having the Intel Core i5-1135G7 CPU and 16 GB RAM.

The entire procedure involved in the wind power prediction is illustrated in Fig. 1. The ML models adopted in this study are presented in the next subsections.

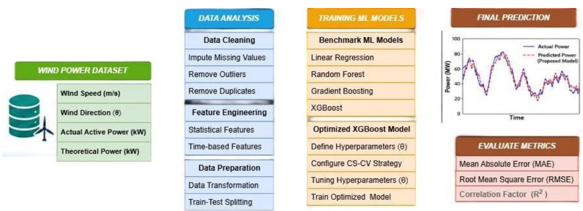

Fig. 1. Overall Process of Wind Power Forecasting

2.1. Data Analysis

During preprocessing, the dataset is examined for missing values, duplicates, outliers, and correlations among attributes. Subsequently, anomalies are identified and removed to produce a clear, appropriate dataset for training the predictive models. Fig. 2(a) and Fig. 2(b) depict the wind power (kW) and wind speed (m/sec) over the entire time series of data, respectively. Fig. 3 shows a heatmap that helps identify correlations between wind active power and other features. Subsequently, a data transformation is applied to normalize feature magnitudes to a comparable range using the Standard Scaler method. During feature engineering, the time column is converted to a date-and-time format. Finally, the dataset is split (train (80%) and test (20%)), and the predictive models are trained accordingly.

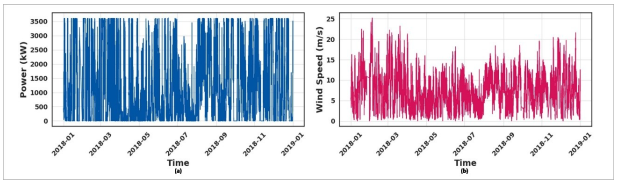

Fig. 2. Wind Power Generation: (a) Wind Power (kW), (b) Wind Speed (m/sec)

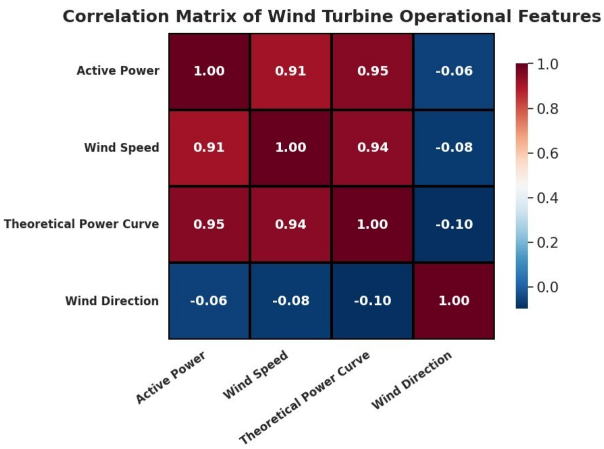

Fig. 3. Correlation Matrix of Heatmap

2.2. Linear Regression (LR)

It determines the linear association between the target variable and input features. The linear regression model is trained on input features, treated as predictors, and a target variable representing wind power output. This predictive method is straightforward and valuable for benchmarking its wind power prediction results against advanced machine learning techniques [9].

For a given dataset, target variable (y), k-length of regressors (x), the expression can be written as [18],

$$\{y_j, x_{j_1}, \ldots, x_{j_k}\}_{j=1}^{n} \tag{1}$$

The mathematical expression of multiple regression can be written as,

$$y_j = \beta_0 + \beta_1 x_{j_1} + \beta_2 x_{j_2} + \ldots + \beta_k x_{j_k} + e_j \tag{2}$$

Where β_0 is the intercept, $(\beta_1 \ldots \beta_k)$ are regression coefficients evaluated from data, and e_j is the residual error.

2.3. Random Forest (RF)

It is an ensemble tree model used for regression and classification applications. By integrating the predictions from multiple decision trees, it enhances the model's stability and decreases volatility [18]. Furthermore, it can handle the high dimensionality of large datasets, making it suitable for predicting wind power, where numerous weather factors influence output. The function of RF predictions is expressed as follows [9]:

$$y_{op} = \frac{1}{n} \sum_{n=1}^{n} P_n(x) \tag{3}$$

Where y_{op} is the predicted output for the assigned input features (x), n stands for number of trees, and $P_n(x)$ is the prediction of n^{th} decision tree.

2.4. Gradient Boosting (GB)

This technique is very helpful for analyzing complex relationships, such as the wind power curve, by applying different loss functions and learning from the input data sample by sample [11]. GB uses gradient-based optimization in regression problems to minimize the loss function. The mathematical representation of GB is given [19] as,

$$y_{op} = F_a(x) = \sum (f_{-}j_{dt}(x)) \tag{4}$$

Where y_{op} denotes output predictions, $F_a(x)$ is the predictions overall, $f_{-}j_{dt}(x)$ is the individual decision tree function with assigned input features (x).

2.5. Extreme Gradient Boosting (XGBoost)

It is potent ensemble technique that reveals hidden patterns in data, leading to more accurate wind power predictions. It consists of an additive regression-tree model, in which every tree is learned to compensate for the residual errors with the current ensemble by utilizing the loss function. Second-order optimization with explicit regularization ensures that XGBoost is both highly accurate and computationally efficient when analyzing large datasets [9]. For this baseline XGBoost model, the major hyperparameters are set to the following default values: estimators = 100, tree depth = 6, and learning rate = 0.3. The formulation of the predictive function is expressed as [16],

Table 2. Results of Predictive Metrics (Training Set)

Models	Train Set			Test Set		
	MAE	RMSE	R^2	MAE	RMSE	R^2
LR	189.62	398.12	0.908	195.20	408.52	0.902
RF	123.72	306.19	0.945	131.50	326.31	0.937
GB	113.22	271.54	0.957	122.25	293.37	0.949
XGBoost	99.89	239.83	0.966	110.59	267.83	0.957
Optimized XGBoost	**56.85**	**119.11**	**0.991**	**81.65**	**201.62**	**0.976**

$$y_j = \sum_{n=1}^{N} (f_n(x_j)) \tag{5}$$

Where y_j is the forecasted output, $f_n(x_j)$ is function of prediction for n^{th} regression tree with given input feature, and N is the number of regression trees, and its objective can be expressed as,

$$\ell = \sum_{j=1}^{k} L(y_j, y_j) + \sum_{n=1}^{N} \rho(f_n) \tag{6}$$

where $L(y_j, y_j)$ denotes the squared error loss function, y_j states predicted output, y_j stands for actual output, and $\rho(f_n)$ gives the term of regularization for tree 'n'.

2.6. Proposed Grid Search Cross Validation (GS-CV) strategy

The improved prediction performance of XGBoost is achieved by tuning its hyperparameters (trees numbers, depth of tree, rate of learning, and sampling ratios) using GSCV strategy. In the optimization process, a finite grid of hyperparameter values is defined, and the model is trained with various combinations of these values. The effectiveness of model is then evaluated by applying cross-validation [16]. The combination of GS-CV helps identify the best hyperparameters and provides unbiased estimates of expected performance on unseen data.

The mathematical expression of GS-CV is expressed as,

$$\theta^* = \min_{\theta \in \Theta} \arg \left(\frac{1}{K} \sum_{k=1}^{k} L_k(\theta) \right) \tag{7}$$

where θ^* denotes the combination of optimal hyperparameters, K states the number of cross validation folds, $L_k(\theta)$ stands for validation loss for k fold with hyperparameters (θ). The process steps of the GS-CV strategy are depicted in Fig. 4 [16], [20]. In this work, the GS-CV method analyzed 72 hyperparameter combinations (216 fits) to find the best hyperparameters of XGBoost method used in wind-power prediction: maximum depth of tree 7; learning rate = 0.07; number of estimators = 300; subsample = 0.8; and Col. sample by tree = 1.0.

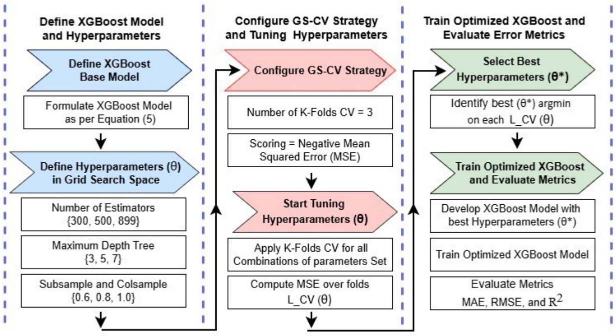

Fig. 4. Process Steps of GS-CV Strategy

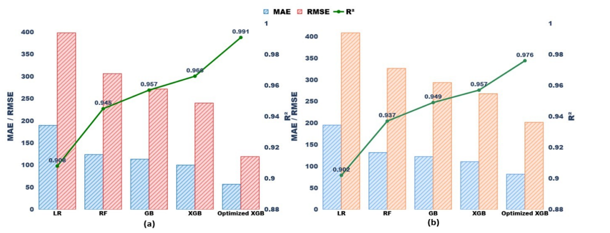

Fig. 5. Predictive Results: (a) Results with Training Set; (b) Results with Test Set

3. Result discussions

The results in Table 2 and Fig. 5 (a) (Training set), and Fig. 5 (b) (Test set) clearly show that the optimised XGBoost model outperforms other ML methods. Compared with other ensemble models, the LR method achieved the highest error metric and lowest R^2 values owing to its insufficient capacity to handle nonlinear interactions. The use of ensemble models RF and GB had a remarkable improvement in the R^2 and reduced error rates, proving their ability to perform well in dealing with nonlinear features of the dataset. XGBoost also outperformed the previous models regarding error prediction. However, the optimized XGBoost model achieved the highest R^2. For test predictions, the optimized XGBoost model reduced MAE and RMSE by approximately 58% and 51%, respectively, compared to LR. The reduction compared to XGBoost is around 26%

and 25%, with a highest R^2 value of 0.976. It is interesting that the training and test R^2 values are close, at 0.991 and 0.976, respectively. The results clearly indicate that the optimized XGBoost model achieves higher prediction accuracy, with significant reductions in error metrics and improvements in R^2 for wind power prediction. These findings have important technical significance, as they indicate that the XGBoost algorithm has become better at capturing the relationship between features and wind power production due to the nonlinearities and non-stationarity present in this map. Poor performance of the linear regression indicates that this approach is not appropriate for this problem, as it cannot capture the complex relationships in the data, whereas the higher performance of RF and GB indicates the importance of applying ensemble learning techniques. In the case of XGBoost, additional benefits arise from sequential boosting, in which residual errors are reduced at each iteration. Finally, optimization contributes to this by adjusting the bias-variance trade-off due to the correct selection of hyperparameters. High similarity of R^2 scores obtained on both training and testing samples indicates good generalization abilities and absence of overfitting, which is critical for wind power forecasting.

The graphs in Fig. 6 (a) and Fig. 6 (b) reveals that the optimal XGBoost model converges faster while providing lower MAE and RMSE values throughout the entire 10,000 samples forecast period. While RF, GB, and base XGBoost perform better compared to LR models, their errors are higher throughout the whole process. It is obvious that GS-CV tuning contributes to more stable predictions.

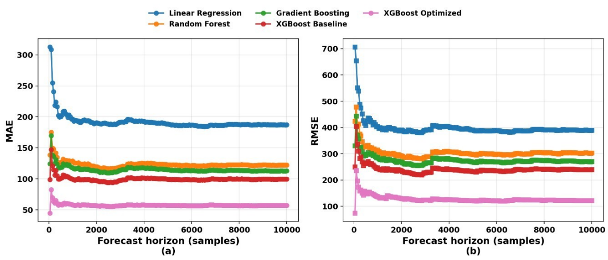

Fig. 6. Error curves over the forecast horizon: (a) MAE; (b) RMSE

Moreover, the regression scatter plots in Fig. 7(a) and Fig. 7(b) show actual versus predicted wind power output for all models, with training- and test-set analyses, respectively. LR (pink) has the largest scatter and strongest bias (~highest power (>1500 kW)) in both graphs, again verifying the poor management of non-linearity, though R^2 = 0.908/0.902 on training and test sets, respectively. RF and GB smoothly improve clustering around the reference line y = x, whereas "regular" XGBoost shows a strong correlation with it.

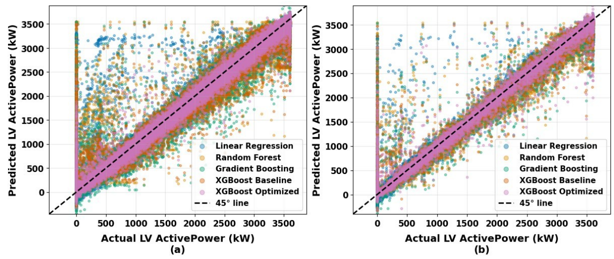

Fig. 7. Regression Plots: (a) Results with training set; (b) Results with testing set

Most importantly, "optimized" XGBoost has the best fit in the range of 0-3000 kW on the training data ($R^2 = 0.991$) and shows the best correlation on the test data too, with only minimal variation $\Delta = 0.015$ and the fewest outliers: an excellent generalization capability, vital for guaranteed wind power prediction. Thus, the optimized XGBoost delivers superior forecasting accuracy across all visualizations, the closest scatter alignment to the perfect prediction line, the flattest error evolution curves, and minimal train-test degradation, demonstrating strong generalization for operational wind farm peak shaving and grid stability.

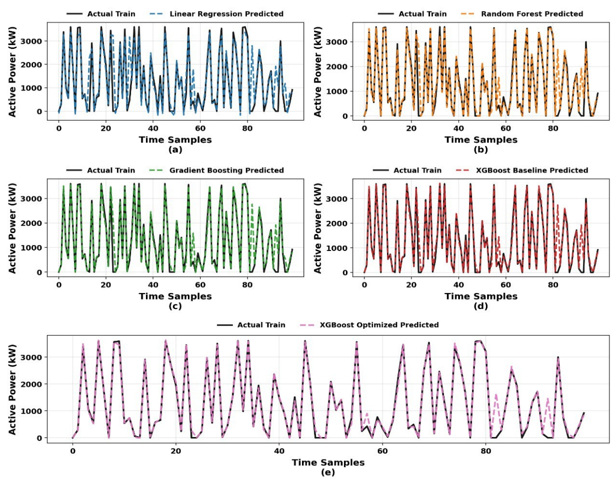

Fig. 8. Time-series prediction across 100 samples horizon: (a) LR; (b) RF; (c) GB; (d) XGBoost; (e) Optimized XGBoost

Fig. 8 shows the time-series predictions (training set) for all models using 100 samples. In Fig. 8(a), the LR model shows a significant phase lag during the ramp stages (20 to 40^{th} samples). There is also an underestimate of values around 20%. The RF model in Fig. 8(b) displays minimal phase lag, but step changes and overshoots are evident at around the 60^{th} sample. The GB model in Fig. 8(c) displays

Table 3. Comparative analysis between the proposed model and existing research works

Reference (Year)	Models	Key Results	Advantages	Limitations	Relevance to present work
[13] (2022)	Bayesian Optimized XGBoost	Improved wind power predictions under transient conditions	Improved predictions with optimized hyperparameters	Metrics are not directly comparable	Supports for non-linear prediction
[9] (2025)	ML, DL, and Ensemble	Stable predictions with higher accuracy	Improved accuracy with ensemble models	Higher computation and complexity	Proves the effectiveness of ensembles
[20] (2024)	DT, RF, KNN, and XGBoost	Higher predictive performance with XGBoost	Good baseline for classical ML models	Results with the Specific dataset	Supports boosted tree learning
[14] (2023)	Feature Engineering with XGBoost	Offers accurate prediction with reduced error metrics	Higher accuracy and low corotational cost	Metrics are not directly comparable	Supports the use of feature engineering with XGBoost
Proposed Model of this study	GS-CV optimized XGBoost	Higher predictive performance on both the train and test set	Higher accuracy and strong generalization	Performance based on the data quality and tuning effort	Represents the robust and practical predictive performance

accurate predictions with minimal errors during the sharp decrease in the 40^{th} to 50^{th} sample zone. The standard XGB model in Fig. 8(d) predicts trends accurately but underestimates values by about 10% during peak transitions. The optimized XGB model in Fig. 8(e) makes accurate predictions that closely match the actual results, especially for wind power. In conclusion, the supremacy of GS-CV-optimized XGBoost model performance on both training and test datasets is unparalleled, with perfect time series fidelity and $R^2 = 0.991$ on the training set and strong test set generalization with $R^2 = 0.976$, minimal $\Delta R^2 = 0.015$, and the tightest scatter to perfect prediction and smoothest error evolution with reductions of $26 - 139\%$ over all other approaches for production-ready wind farm dispatch and peak shaving capabilities of $20 - 30\%$. In summary, the optimized hyperparameters of the proposed XGBoost model offer superior predictive performance to other models. The proposed model is robust in representing the nonlinear input-output relationship with improved generalization performance. The proposed model can be helpful for operational wind forecasting due to its high performance in terms of accuracy and generalization. This performance can be used for grid operations, dispatching, reserves, and wind energy integration.

In addition, a comparison between the proposed GS-CV optimized XGBoost model and recent research works in wind power forecasting is presented in Table 3. It can be noted from the literature that the optimized boosting algorithms and ensemble learning, along with the feature-engineered XGBoost models, yield good performance, especially for nonlinear dynamics of wind power. The proposed model demonstrates high accuracy for both training and testing sets, suggesting good generalization and robust forecasting behavior. On the other hand, the comparison demonstrates that performance is still highly dependent on data and feature quality, which illustrates the advantage and limitation of the proposed approach.

4. Conclusions

Overall, GS-CV-optimized XGBoost is a state-of-the-art benchmark model for providing accurate wind power predictions. The optimized XGBoost model achieves higher R^2 values on both the training set ($R^2 = 0.991$) and the test set ($R^2 = 0.976$). Compared with other models (LR, RF, GB, and standard XGBoost), the optimized XGBoost model achieved error reductions of 26% to 139%. Also, the optimized XGBoost model provides firm regression alignment with the reference line's actual values and maintains a close relationship, with less variation between actual and predicted wind power over the horizon. The results clearly demonstrate that the GS-CV-optimized XGBoost model is more robust, with improved accuracy and generalization in wind power forecasts, providing a practical, data-driven solution for power system operators. Nevertheless, the study has some limitations due to the use of only one year of data from a single location, which could limit the applicability of the results to different wind scenarios. Therefore, future studies will have to focus on validating the suggested methodology using data from multiple locations, adding other meteorological parameters, and comparing its performance with that of state-of-the-art deep learning approaches.

Acknowledgements

The authors would like to thank New Horizon College of Engineering for providing facilities and kind support to carry out this research.

Authors' contributions

All authors contributed to the study conception and design, writing, and editing: Writing and model development: A.V; Model analysis: KMA and DA; Editing and final validation: SKH and AS.

Funding

No external funding was received for this research

Data availability

This research used a real-time wind power dataset from the Kaggle database, available via the link in the References section.

Declarations

Competing interests

The authors declare no competing interests.

References

[1] E. A. Tuncar, Ş. Sağlam, and B. Oral, (2024) *"A Review of Short-Term Wind Power Generation Forecasting Methods in Recent Technological Trends"* **Energy Reports *12***: 197–209. DOI: 10.1016/j.egyr.2024.01.12.3.

[2] H. R. Alsamamra, S. Salah, and J. H. Shoqeir, (2024) *"Performance Analysis of ARIMA Model for Wind Speed Forecasting in Jerusalem, Palestine"* **Energy Exploration & Exploitation *42***(5): 1727–1746. DOI: 10.1177/01445987241234567.

[3] A. Alkesaiberi, F. Harrou, and Y. Sun, (2022) *"Efficient Wind Power Prediction Using Machine Learning Methods: A Comparative Study"* **Energies *15***(7): 2327. DOI: 10.3390/en15072327.

[4] Z.-Y. Fu, G.-Q. Li, W.-D. Tang, Z.-H. Pan, and W.-C. Zhang, (2025) *"Novel Nonlinear Wind Power Prediction Based on Improved Iterative Algorithm"* **Systems Science & Control Engineering *13***(1): 2448626. DOI: 10.1080/21642583.2025.2448626.

[5] Z. Ti, X. W. Deng, and M. Zhang, (2021) *"Artificial Neural Networks Based Wake Model for Power Prediction of Wind Farm"* **Renewable Energy *172***: 618–631. DOI: 10.1016/j.renene.2021.03.056.

[6] D. A. Savio, V. A. Juliet, B. Chokkalingam, S. Padmanaban, J. B. Holm-Nielsen, and F. Blaabjerg, (2019) *"Photovoltaic Integrated Hybrid Microgrid Structured Electric Vehicle Charging Station and Its Energy Management Approach"* **Energies *12***(1): 168. DOI: 10.3390/en12010168.

[7] J. Ferrero Bermejo, J. F. Gomez Fernandez, F. Olivencia Polo, and A. Crespo Márquez, (2019) *"A Review of the Use of Artificial Neural Network Models for Energy and Reliability Prediction: Solar PV, Hydraulic and Wind Energy Sources"* **Applied Sciences *9***(9): 1844. DOI: 10.3390/app9091844.

[8] Z. Liu, H. Guo, Y. Zhang, and Z. Zuo, (2025) *"A Comprehensive Review of Wind Power Prediction Based on Machine Learning: Models, Applications, and Challenges"* **Energies *18***(2): 350. DOI: 10.3390/en18020350.

[9] Z. Mustaffa and M. H. Sulaiman, (2025) *"Random Forest Based Wind Power Prediction Method for Sustainable Energy System"* **Cleaner Energy Systems *12***: 100210. DOI: 10.1016/j.cles.2025.100210.

[10] T. A. Rajaperumal and C. C. Columbus, (2025) *"Enhanced Wind Power Forecasting Using Machine Learning, Deep Learning Models and Ensemble Integration"* **Scientific Reports *15***(1): 20572. DOI: 10.1038/s41598-025-20572-3.

[11] R. A. Sobolewski, M. Tchakorom, and R. Couturier, (2023) *"Gradient boosting-based approach for short-and medium-term wind turbine output power prediction"* **Renewable Energy *203***: 142–160. DOI: 10.1016/j.renene.2022.12.040.

[12] S. Ponrekha A, M. S. P. Subathra, C. Bharatiraja, and N. Manoj Kumar, (2025) *"A Topology Review and Comparative Analysis on Transformerless Grid-Connected Photovoltaic Inverters and Leakage Current Reduction Techniques"* **IET Renewable Power Generation *19***(1): e12655. DOI: 10.1049/rpg2.12655.

[13] X. Xiong, X. Guo, P. Zeng, R. Zou, and X. Wang, (2022) *"A Short-Term Wind Power Forecast Method via XGBoost Hyper-Parameters Optimization"* **Frontiers in Energy Research *10***: 905155. DOI: 10.3389/fenrg.2022.905155.

[14] S. Guan, Y. Wang, L. Liu, J. Gao, Z. Xu, and S. Kan, (2023) *"Ultra-Short-Term Wind Power Prediction Method Combining Financial Technology Feature Engineering and XGBoost Algorithm"* **Heliyon *9***(6): e16938. DOI: 10.1016/j.heliyon.2023.e16938.

[15] Xiangcheng, J. Wang, Z. Geng, Y. Jin, and J. Xu, (2023) *"Short-Term Wind Power Prediction Method Based on Genetic Algorithm Optimized XGBoost Regression Model"* **Journal of Physics: Conference Series *2527***(1): 012061. DOI: 10.1088/1742-6596/2527/1/012061.

[16] N. T. Tran, T. T. G. Tran, T. A. Nguyen, and M. B. Lam, (2023) *"A new grid search algorithm based on XGBoost model for load forecasting"* **Bulletin of Electrical Engineering and Informatics *12***(4): 1857–1866. DOI: 10.11591/eei.v12i4.4907.

[17] B. İsen. *Wind Turbine SCADA Dataset*. Accessed: February 13, 2026. 2018.

[18] C. El-Morr, M. Jammal, H. Ali-Hassan, and W. El-Hallak, (2022) *"Machine Learning for Practical Decision Making"* **International Series in Operations Research and Management Science**: DOI: https://doi.org/10.1007/978-3-031-16990-8.

[19] U. Singh, M. Rizwan, M. Alaraj, and I. Alsaidan, (2021) *"A Machine Learning-Based Gradient Boosting Regression Approach for Wind Power Production Forecasting: A Step Towards Smart Grid Environments"* **Energies *14***(16): 5196. DOI: 10.3390/en14165196.

[20] A. Bilgili and K. Gül, (2024) *"Forecasting power generation of wind turbine with real-time data using machine learning algorithms"* **Clean Technologies and Recycling *4***: 108–124. DOI: 10.3934/ctr.2024006.