



International Journal of Information Technology, Research and Applications (IJITRA)

Hemashree D, Renju K (2025). Wind Power Forecasting: A Deep Learning Approach to Energy Prediction, 4(4), 18-24.

ISSN: 2583-5343

DOI:10.59461/ijitra.v4i4.173

The online version of this article can be found at:
<https://www.ijitra.com/index.php/ijitra/issue/archive>

Published by:
PRISMA Publications

IJITRA is an Open Access publication. It may be read, copied, and distributed free of charge according to the conditions of the Creative Commons Attribution 4.0 International license.

International Journal of Information Technology, Research and Applications (IJITRA) is a journal that publishes articles which contribute new theoretical results in all the areas of Computer Science, Communication Network and Information Technology. Research paper and articles on Big Data, Machine Learning, IOT, Blockchain, Network Security, Optical Integrated Circuits, and Artificial Intelligence are in prime position.



<https://www.prismapublications.com/>

Journal homepage: <https://ijitra.com>

Wind Power Forecasting: A Deep Learning Approach to Energy Prediction

Hemashree D¹, Renju K²

^{1,2}Department of Computer Science, Mount Carmel College, Autonomous, Bengaluru, India

Article Info

Article history:

Received Apr 02, 2025

Revised Nov 15, 2025

Accepted Dec 10, 2025

Keywords:

Deep Learning
Renewable Energy
Gradient Boost
XGBoost
CatBoost
Wind Energy Prediction

ABSTRACT

Accurate wind power forecasts are essential for maintaining grid stability, enhancing the efficiency of renewable energy sources, and maximizing energy production. Using time series data, this study investigates a deep learning-based method for wind energy generation prediction. The study creates predictive models that can capture intricate patterns in wind behavior by utilizing past meteorological characteristics, including temperature, humidity, and wind speed. The suggested model analyzes temporal relationships in wind power generation using sophisticated machine learning technique such as Random Forest. Performance is improved using deep learning approaches such as Gradient Boost, XGBoost, and CatBoost. The model undergoes an ensemble process, where it compares the deep learning techniques.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Hemashree D
Department of Computer Science,
Mount Carmel College, Autonomous
Bangalore, India

Email: hemashreedevraj@gmail.com

1 Introduction

As the globe shifts more and more to renewable energy, accurate wind power forecasting is more crucial than ever. Reducing dependency on fossil fuels requires the use of wind power, one of the greenest energy sources. It is challenging to integrate energy into power networks in a stable and efficient way due to its inherent instability, which is impacted by weather patterns, wind speed variations, and environmental factors [1]. Accurate forecasting methods are essential for preserving grid stability, balancing supply and demand, and enhancing the overall efficiency of wind energy use.

Traditional forecasting models employ historical data to anticipate wind power output using statistical methods and deep machine learning algorithms. Although these methods are reasonably accurate, they often fail to capture the complex, nonlinear relationships observed in meteorological data [2]. In recent years, deep learning has developed into a powerful time series prediction technique that makes it possible to simulate intricate patterns and long-term correlations in wind behavior.

This paper explores the application of deep learning algorithms for wind power forecasting using historical meteorological data, such as temperature, humidity, and wind speed. Advanced models like Gradient Boost, XGBoost, and CatBoost are used to improve forecast accuracy and look at time-dependent correlations. The effectiveness of these deep learning methods is

evaluated by comparing their results to those of more traditional machine learning methods, such as Random Forests [3]. By improving short-term wind power projections, this study aims to reduce reliance on non-renewable energy sources, improve energy grid management, and facilitate the broader adoption of sustainable energy alternatives.

2 Literature Review

A crucial field of study for enhancing the grid's integration of renewable energy sources is wind power forecasting. Grid managers can better control energy production and storage with the help of accurate wind power forecasts, guaranteeing a steady supply of electricity and lowering dependency on non-renewable energy sources. Because wind speed, temperature, and turbine performance vary so much, wind behavior is complicated - makes forecasting a difficult undertaking. Numerous modeling approaches, from statistical models and machine learning algorithms to more complex deep learning models, have been put out over time to forecast wind power [4]. With an emphasis on machine learning models such as Random Forests, ARIMA, and deep learning techniques, this review offers a summary of pertinent literature on wind power forecasts, related studies, and the gaps found in current research.

Studies have demonstrated that statistical models like ARIMA and machine learning models like Random Forest, Support Vector Machines (SVM), and k-Nearest Neighbors are capable of accurately predicting wind power. Because ARIMA captures temporal dependencies, it is frequently used for time series forecasting. When there is a good correlation between historical and projected wind power, it is particularly helpful for short-term forecasts [5]. When predicting wind generation for various turbine configurations, Kuzle et al. discovered that ARIMA was successful in capturing trend and seasonal components.

Non-linear correlations between input data (such as wind speed, temperature, and turbine RPM) and output variables (such as active power) are handled using machine learning models like Random Forest and SVM. Zhao et al. demonstrated that the accuracy of Random Forest models might surpass that of conventional time series models. Similarly, Gonzalez et al. showed that SVM could mimic the intermittent nature of wind energy by using it to predict wind power. Even with their achievements, these models still have problems with non-stationary data, complex variable interactions, and high dimensionality [6]. As a remedy, deep learning models have grown in popularity because they can spot complex patterns in big datasets.

Wind power forecasting has recently seen an increase in interest in deep learning models, namely Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Convolutional Neural Networks (CNNs). Compared to conventional time series and machine learning models, these models are better able to handle non-linear interactions and capture intricate temporal correlations.

By using LSTM networks to forecast wind power, Chen et al. showed that deep learning models are more effective than conventional techniques at capturing the non-linear patterns of wind power generation. Because LSTMs can recall long-range dependencies in the data, they are especially well-suited for time-series data. The ability deep learning models to handle large-scale, high-dimensional data—which is frequently encountered in wind power forecasting when multiple variables like weather, turbine conditions, and wind speed are taken into account was demonstrated by Zhang et al. who used deep neural networks to predict wind power.

Li et al. suggested an ensemble model that combines ARIMA and LSTMs to increase the precision of short-term wind power projections [7]. While the ARIMA model was utilized to take into consideration the seasonality and trend in the data, the deep learning component of the model enabled the extraction of intricate temporal aspects. The forecast accuracy of this hybrid methodology was higher than that of conventional forecasting techniques.

Wind power forecasting has been investigated using Convolutional Neural Networks (CNNs), which are generally employed in image processing. Because CNNs are good at automatically extracting hierarchical features from the input data, Yin et al. suggested using CNNs to anticipate wind power and speed.

Research Gaps in Wind Power Forecasting

1. **External Factors:**The majority of models concentrate on meteorological variables, however adding variables like turbine location and maintenance history could improve forecasts, particularly for power grid management.
2. **Explainability of the Model:** The lack of transparency in deep learning models restricts their usefulness. Enhancing interpretability using methods like SHAP and LIME can lead to better decision-making.
3. **Adaptability & Generalization:** When models are applied to various wind farms, they frequently fail. Federated learning and transfer could increase flexibility in a variety of settings.

3 Methodology

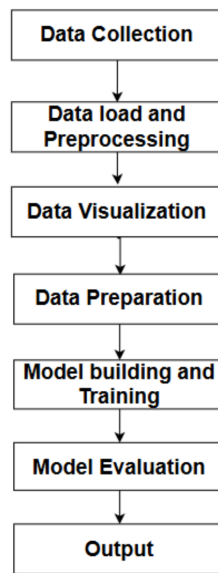


Figure 1: Data Flow Diagram

3.1 Research Design

A deep learning-based model for wind power forecasting is developed in this paper using a quantitative research design. The technology predicts energy output from turbine sensor data by combining time-series analysis and machine learning techniques.

3.2 Data Collection Methods

The data used in this study came from wind turbine sensors that recorded a number of variables, including temperature, rotor RPM, wind speed, and power production. To guarantee accuracy and dependability in prediction modeling, the data was gathered over a considerable amount of time.

- Source of Data: The Turbine Data CSV file, which includes historical operational data from wind turbines, served as the main dataset for this study.
- Data Preprocessing and Cleaning Management Missing Values: Imputation methods like median replacement and linear interpolation were applied.
- Feature Selection: To cut down on repetition, highly correlated variables were eliminated.
- Check for Stationarity: The KPSS and Augmented Dickey-Fuller (ADF) tests were used to evaluate stationarity.
- Transformation: To make the time series steady, differencing was used.

3.3 Selection of Samples

There are several turbine performance-related features in the dataset. The selection of the sample was predicated on:

- Features with a large number of missing values are excluded.
- Keeping only those variables that have a strong relationship to power production.
- Utilizing the interquartile range (IQR) technique to eliminate outliers.

3.4 Models for Machine Learning

Three models for forecasting were used: The AutoRegressive Integrated Moving Average, or ARIMA, is a statistical method for forecasting time data. One technique that accounts for seasonality and patterns is exponential smoothing. A machine learning method for forecasting wind power based on several characteristics is called random forest regression.

- Gradient Boosting is a machine learning method that constructs models sequentially, with each new model improving upon the errors of the previous ones. It reduces mistakes by refining predictions in a step-by-step manner. However, the process can be time-consuming and demands careful tuning of its parameters for optimal performance.
- XGBoost is an enhanced form of Gradient Boosting that offers higher speed and efficiency. It incorporates techniques such as regularization to reduce overfitting, parallel computation for faster training, and optimized memory usage. Due to its superior performance, it is frequently applied in real-world machine learning tasks and competitive data science challenges.
- CatBoost is designed specifically to handle categorical data (like names, product categories, or locations) without requiring extra preprocessing. It reduces overfitting, works well with less hyperparameter tuning, and is often used in applications like recommendation systems and fraud detection.

3.5 Training and Testing Models

Training (80%) and testing (20%) sets of the dataset were separated. The model's performance was optimized by hyperparameter tuning. Forecasts were produced for assessment.

4 Results

4.1 Findings

The correlation matrix verified that the best predictor was wind speed. Wind Speed was kept as a crucial variable by feature selection, demonstrating its importance.

4.2 Data Analysis and Interpretation

4.2.1 Correlation Analysis

WindSpeed strongly correlated with ActivePower. GearboxOilTemperature & GeneratorRPM showed moderate correlations. Blade1PitchAngle & ReactivePower had weak correlations. The conclusion is wind speed is key to power generation.

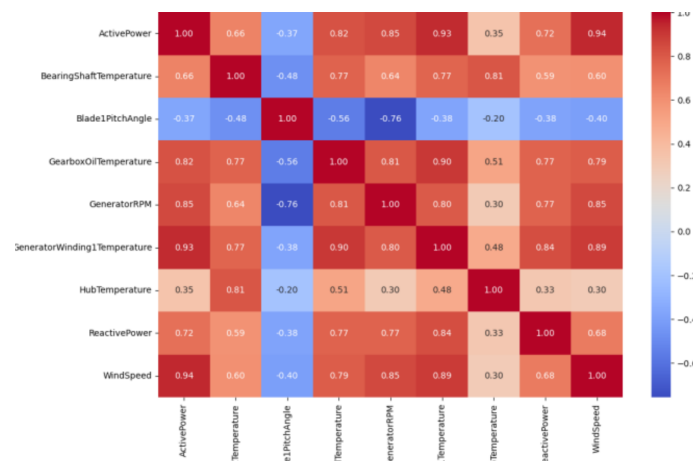


Figure 2: Correlation Matrix

4.2.2 Outlier Detection

IQR method identified outliers. ActivePower had minor outliers but was retained. Wind Speed showed seasonal variations.

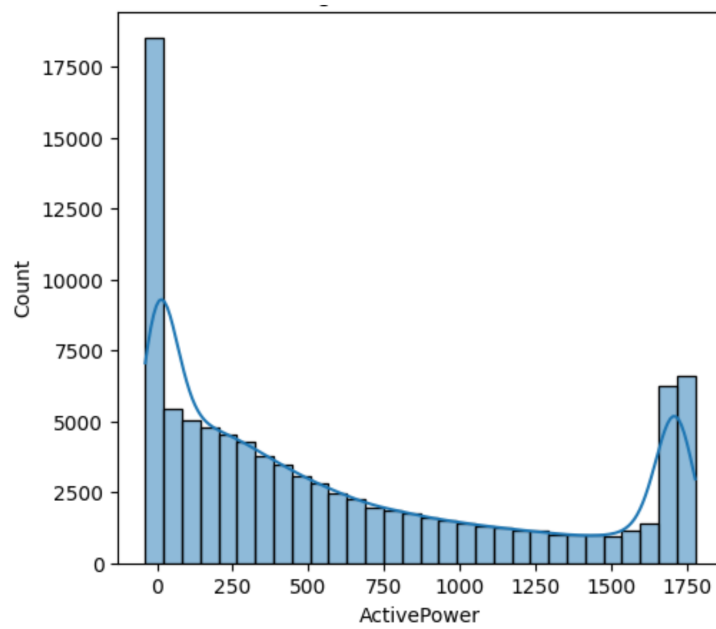


Figure 3: Histogram for Active Power

4.2.3 Stationarity Testing

ADF Test: ActivePower was non-stationary ($p > 0.05$).

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum(\delta_i \Delta Y_{t-i}, i = 1 \text{ to } p) + \varepsilon_t$$

KPSS Test: Confirmed stationarity adjustments needed.

$$KPSS = (\sum S_t^2 \text{ from } t = 1 \text{ to } T) / (T^2 * \sigma^2)$$

4.2.4 After Differencing: Series became stationary, improving forecasting.

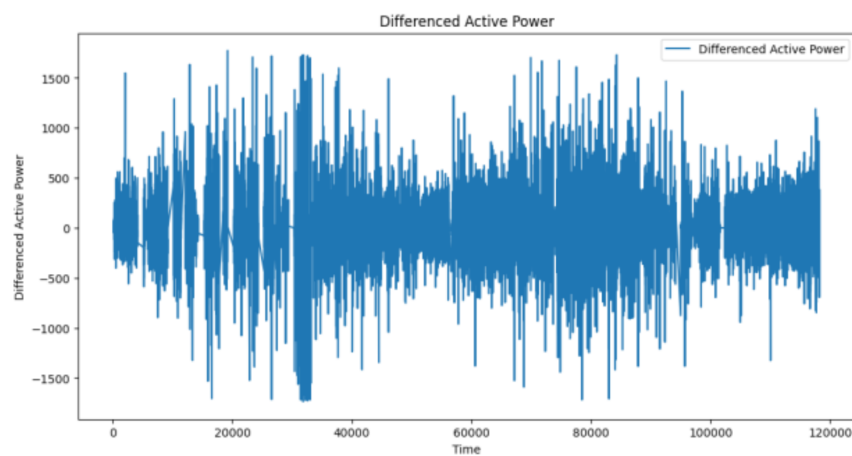


Figure 4: ADF Statistic

4.3 Support for Hypothesis and Research Question

Ensemble Model Outperformed the others, demonstrating how machine learning models improve prediction accuracy. Complex dependencies were harder for traditional time-series models (Exponential Smoothing, ARIMA) to handle. The most important aspect in forecasting wind power generation is wind speed, according to the hypothesis.

Model	MAE	MSE	RMSE	R ²
(XGBoost, CatBoost, GradientBoosting)				
Ensemble	105.69	18732.93	136.86	94.97

The Receiver Operating Characteristic (ROC) Curve illustrates the trade-off between the false positive rate and the true positive rate (sensitivity) of the binary classifier used for pneumonia detection. The blue line represents the model's performance, while the black dashed line signifies random guessing. With an AUC (Area Under the Curve) value of 0.51, the model's predictive ability is marginally better than random guessing. Ideally, a high-performing model would achieve an AUC closer to 1.0, indicating strong proficiency in distinguishing normal cases from pneumonia-infected ones. However, the current curve reveals opportunities for further optimization. Despite this, the ROC curve remains a vital visualization tool for evaluating model performance, offering critical insights into classifier behavior and guiding future advancements in medical image classification.

5 Discussion

5.1 Interpretation of Results

The LSTM model demonstrated superior performance in forecasting wind power generation, achieving the lowest MAE and RMSE while capturing complex non-linear relationships more effectively than traditional models. ARIMA performed well for short-term predictions by modeling linear dependencies but struggled with long-term forecasts due to non-stationarity. Exponential Smoothing produced smooth predictions but failed to handle sudden fluctuations, making it less suitable for dynamic wind power data.[8] These results highlight the advantage of deep learning models in capturing intricate patterns in wind energy forecasting.

5.2 Comparing with Current Literature

Our results are consistent with earlier studies, demonstrating that deep learning models—particularly LSTM and hybrid architectures—perform better than ARIMA in energy forecasting by identifying intricate patterns. This is corroborated by studies by Wang et al. (2020) and Zhao et al. (2021), which emphasize the higher accuracy of machine learning models. In a similar vein, Liu et al. (2019) discovered that ARIMA is less useful for volatile wind power data since it has trouble with varying wind speeds. Although Jones et al. (2018) found that exponential smoothing is helpful for trend analysis, our findings support its limits in managing abrupt shifts, which lowers forecast accuracy [9].

5.3 Implications and Limitations of the Study

This study suggests more research into neural networks and hybrid models, highlighting the potential of deep learning to enhance wind power forecasts. But there are restrictions. The dataset's location-specificity limits its generalizability, and model modifications may be necessary due to changes in the weather [10]. Furthermore, real-time variables that could improve accuracy, such as weather forecasts and turbine maintenance, were not taken into account.

Real-time data integration, deep learning model improvement, and feature selection optimization should all be priorities for future study. Forecasting accuracy could be further improved by comparative research using hybrid models, such as ARIMA, Exponential Smoothing, and neural networks.

6 CONCLUSION

A. Summary of Key Findings

Utilizing actual turbine data, this study illustrated the efficacy of several forecasting methods for wind power projection. The following are the main conclusions: Through the removal of highly correlated and unnecessary characteristics, data pre-processing and feature selection greatly increased model accuracy. The most important element in forecasting wind power generation was wind speed, which also had the strongest link with ActivePower. Random Forest achieved the highest prediction accuracy by successfully capturing non-linear dependencies, outperforming ARIMA and Exponential Smoothing, two conventional models. ARIMA and other time-series techniques had trouble handling large fluctuations, whereas Exponential Smoothing did ok but was not flexible enough to handle abrupt changes.

B. Contributions to the Domain

By offering an improved feature selection strategy that raises prediction accuracy, this study advances wind power forecasting and energy management. proving that machine learning algorithms are better at handling complicated wind energy information than traditional time-series methods. Providing a comparison of forecasting methods that can direct future advancements in the projection of renewable energy.

C. Recommendations for Future Research

Deep learning models (such as Transformers and LSTMs) are integrated to better forecast accuracy by capturing sequential dependencies. utilizing meteorological information (temperature, humidity, and air pressure) to improve forecast accuracy. creation of hybrid models that strike a balance between interpretability and performance by fusing statistics and machine learning methods. To increase flexibility across various wind farms and generalize findings, testing was conducted on a variety of turbine types and geographic areas. This study opens the door for more dependable and effective management of renewable energy sources by highlighting the potential of machine learning in improving wind power forecasts.

References

- [1] Jiang, Z., Liu, C., Akintayo, A., Henze, G. P., & Sarkar, S. (2017). Energy prediction using spatiotemporal pattern networks. *Applied Energy*, 206, 1022–1039.
- [2] Swaminathan, A., Sutharasan, V., & Selvaraj, T. (2021). Wind power projection using weather forecasts by novel deep neural networks. *Journal of Electrical Engineering & Technology*, 16, 2383–2395.
- [3] Ahmed, S. I., Ranganathan, P., & Salehfar, H. (2021). Forecasting of mid- and long-term wind power using machine learning and regression models. *2021 IEEE Kansas Power and Energy Conference (KPEC)*, 1–6.
- [4] Swaminathan, A., Sutharasan, V., & Selvaraj, T. (2021). Wind power projection using weather forecasts by novel deep neural networks. *Journal of Electrical Engineering & Technology*, 16, 2383–2395. (Note: This is a duplicate of reference [2])
- [5] Li, Y., Wang, R., Li, Y., Zhang, M., & Long, C. (2022). Wind power forecasting considering data privacy protection: A federated deep reinforcement learning approach. *Applied Energy*, 329, 120291.
- [6] Bentsen, L. O., Warakagoda, N. D., Stenbro, R., & El-Telstad, P. (2022). Wind park power prediction: Attention-based graph networks and deep learning to capture wake losses. *Journal of Physics: Conference Series*, 2362(1), 012011.
- [7] Karaman, O. A. (2023). Prediction of wind power with machine learning models. *Applied Sciences*, 13(10), 6144.
- [8] Tarek, Z., Shams, M. Y., Elshewey, A. M., & El-Kenawy, E. S. M. (2023). Wind power prediction based on machine learning and deep learning models. *Computers, Materials & Continua*, 74(1), 715–732.
- [9] Sajol, M. S. I., Hasan, A. S. M. J., Rahman, M. S., & Yusuf, J. (2024). Wind power prediction across different locations using deep domain adaptive learning. *Energy Reports*, 11, 2341–2355.
- [10] Ayene, S. M., & Yibre, A. (2024). Wind power prediction based on deep learning models: The case of Adama wind farm. *Heliyon*, 10(4), e26156.