



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 15, Issue 3, March 2026**

# **Weather-Based Prediction of wind Turbine Energy Output: A Next-Generation Approach to Renewable Energy Management**

**Bollepalli Teja<sup>1</sup>, Bathula Krishna Priyanka<sup>2</sup>, Bandi Pavan Chaitanya<sup>3</sup>, John Blesson<sup>4</sup>,  
E. Sambasivarao<sup>5</sup>**

UG Students, Department of CSE (AI & ML), Kallam Haranadhareddy Institute of Technology, Guntur, India<sup>1-4</sup>

Associate Professor, Department of CSE (AI & ML), Kallam Haranadhareddy Institute of Technology, Guntur, India<sup>5</sup>

**ABSTRACT:** The growing global transition toward renewable energy has increased the demand for intelligent, data-driven technologies that enhance the reliability and efficiency of power generation systems. Wind energy, one of the most widely utilized renewable sources, is strongly influenced by dynamic and unpredictable meteorological conditions. Conventional prediction techniques often struggle to capture the complex nonlinear relationships among factors such as wind speed, humidity, temperature, and atmospheric pressure, leading to inconsistent forecasts and difficulties in maintaining grid stability. To overcome these limitations, this study presents a machine learning-based predictive model, focusing on the Random Forest algorithm, to accurately estimate wind turbine energy output using weather-related data. By analysing key environmental parameters, the system delivers a more reliable and consistent approach for forecasting energy production.

**KEYWORDS:** Wind Energy Prediction, Machine Learning, Random Forest, Meteorological Data, Renewable Energy Forecasting, Energy Output Estimation, API Integration

## **I. INTRODUCTION**

Wind Turbine Energy Prediction involves estimating the electrical power output generated by a wind turbine by analysing various atmospheric and environmental conditions. These conditions primarily include wind speed, air pressure, temperature, humidity, and wind direction. Since wind is an inherently fluctuating and unpredictable resource, achieving accurate predictions of its behaviour is essential for maximizing energy generation efficiency and minimizing potential losses. Forecasting systems for wind energy rely extensively on large volumes of data gathered from sources such as weather monitoring stations, sensing devices, and satellite observations. This data is systematically analysed to uncover patterns and relationships between changing weather conditions and turbine performance, which form the basis for reliable and effective energy prediction.

Machine learning has emerged as a powerful approach for significantly enhancing the precision of wind energy forecasting systems. Unlike conventional statistical models, machine learning techniques—including Random Forest, Support Vector Machines (SVM), and Neural Networks—are capable of capturing complex and nonlinear dependencies present within large-scale datasets. By learning from historical meteorological data, these models can generate accurate predictions of future power output even under uncertain and rapidly changing environmental conditions. Such predictive insights are highly valuable for wind farm operators and energy providers, as they support informed decision-making related to turbine control, maintenance scheduling, and efficient energy distribution, ultimately improving both operational reliability and economic efficiency. In addition, precise wind energy prediction plays a crucial role in maintaining the stability of modern power grids and accelerating the development of renewable energy technologies. It facilitates the effective integration of wind power into the overall energy mix, reduces dependence on fossil fuel-based generation, and contributes to the reduction of carbon emissions. With continuous advancements in data acquisition technologies, cloud-based computing platforms, and AI-driven analytical methods, prediction systems are becoming increasingly robust, scalable, and capable of delivering real-time results. Therefore, wind turbine energy prediction not only strengthens sustainable energy management practices but also contributes to the development of intelligent, data-centric power systems designed to meet future global energy requirements.

## II. LITERATURE SURVEY

Wind energy forecasting has emerged as a significant research domain within the fields of renewable energy and artificial intelligence, primarily driven by the increasing global emphasis on sustainable and efficient power generation. In earlier stages, research in this area predominantly relied on statistical approaches such as the Weibull distribution to study wind behaviour and turbine performance. Although these methods established a fundamental understanding, they were limited in effectively modelling the complex, nonlinear, and time-varying interactions among meteorological variables, which often resulted in reduced prediction accuracy under fluctuating environmental conditions.

With the rapid evolution of machine learning (ML) methodologies, research has progressively shifted toward data-driven techniques capable of handling large-scale and complex datasets while identifying hidden patterns within environmental features. Kusiak et al. (2009) demonstrated that data mining-based approaches significantly enhance the accuracy and dependability of wind energy predictions when compared to conventional statistical models. Likewise, Li et al. (2021) presented a detailed survey of machine learning applications in wind power forecasting, emphasizing the effectiveness of ensemble-based models such as the Random Forest (RF) algorithm in improving prediction stability and performance [3]. Among the various machine learning techniques, Random Forest has gained considerable attention due to its robustness, flexibility, and strong resistance to overfitting, achieved by aggregating multiple decision trees to produce stable and reliable predictions. Zhang, Wang, and Li (2020) further confirmed that Random Forest models surpass traditional single-model methods by delivering higher accuracy, consistency, and reliability in wind energy forecasting tasks.

### 1) FLASK

Flask is an open-source web framework developed in Python, widely used for building efficient, scalable, and dynamic web applications. It is classified as a micro-framework because it provides only the essential components required for web development without enforcing strict design patterns or unnecessary complexity. This lightweight architecture allows developers the flexibility to integrate only the required tools and extensions based on specific project needs. Due to its simplicity and adaptability, Flask is particularly well-suited for deploying machine learning models and developing interactive platforms for real-time data analysis and prediction.

### 2) JAVA SCRIPT

JavaScript is a highly versatile and widely used programming language defined by the ECMAScript specification (ECMA-262, 14th Edition, 2023). It plays a crucial role in modern web development by enabling dynamic behaviour, user interaction, and real-time responsiveness within web applications. Along with HTML and CSS, JavaScript forms one of the three core technologies of the web, allowing developers to build interactive, responsive, and user-centric applications. Its flexibility enables usage across both client-side and server-side environments, facilitating seamless integration with various frameworks, libraries, and APIs to develop complete full-stack solutions.

### 3) RANDOM FOREST

The Random Forest algorithm is an ensemble-based machine learning technique that combines multiple decision trees to enhance prediction accuracy and model stability. During the training phase, the algorithm constructs numerous decision trees using different subsets of the data and aggregates their outputs to produce a final prediction. This ensemble strategy helps in reducing overfitting and improving generalization, especially when dealing with complex and high-dimensional datasets. In wind energy forecasting, Random Forest has proven to be highly effective in analysing meteorological variables such as wind speed, temperature, humidity, and air pressure.

### 4) PICKLE

The Pickle module in Python is a built-in library used for serializing and deserializing Python objects, allowing complex data structures, trained models, and configurations to be stored and retrieved efficiently. The process of serialization (pickling) converts Python objects into a byte stream, which can be easily saved to a file or transmitted across systems. Deserialization (unpickling) reconstructs the byte stream back into its original object form.

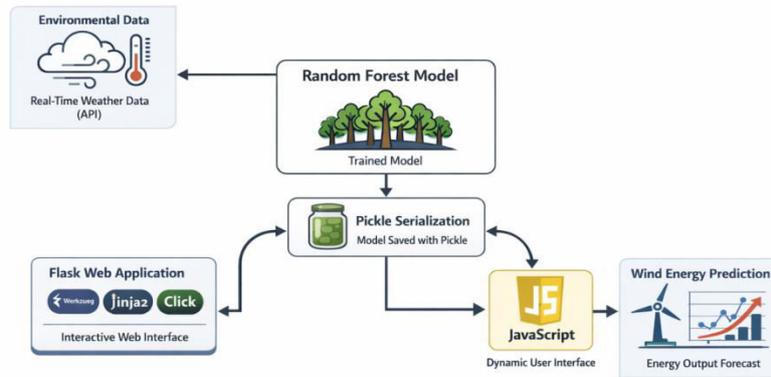


Fig 1:ML System Workflow

### III. PROPOSED METHODOLOGY

The proposed methodology is centred on designing and implementing a machine learning–driven predictive system to estimate wind turbine energy output by utilizing key meteorological variables such as wind speed, temperature, humidity, and air pressure.

Multiple machine learning models, such as Linear Regression, Random Forest, and Neural Networks, are developed and comparatively analysed. Their performance is evaluated using standard metrics including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination ( $R^2$ ). Based on this evaluation, the Random Forest algorithm is chosen due to its high accuracy, robustness, and effectiveness in handling high-dimensional data while minimizing the risk of overfitting.

Following model selection, the Random Forest model is trained and fine-tuned through hyperparameter optimization to improve its generalization capability. The optimized model is then validated using unseen test data to ensure its predictive reliability. After successful validation, the trained model is serialized using the Pickle library in Python, allowing efficient storage and reuse without retraining. For real-world implementation, a web-based application is developed, where the backend is built using Flask and the frontend is designed using HTML and CSS. This interface enables users to input meteorological parameters and receives real-time predictions of wind energy output. Additionally, the system is designed to support integration with live weather APIs, enabling dynamic and real-time forecasting capabilities.



Fig 2: Wind Turbine Energy Prediction Workflow

#### IV. EXPERIMENTAL RESULTS

##### 1. Experimental Configuration

The Wind Turbine Energy Prediction System was evaluated using a dataset comprising key meteorological attributes such as wind speed, temperature, humidity, and air density, along with corresponding power output values. The dataset was partitioned into training and testing sets in an 80:20 ratio. Prior to modelling, feature scaling was applied through standardization, and extreme values were controlled using Historization within the 2nd–98th percentile range. Model optimization was carried out using 5-fold cross-validation on the training data. Performance was assessed using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination ( $R^2$ ).

Hardware/Software Environment: Python 3.11, scikit-learn, Flask; Intel i5 processor with 16 GB RAM (CPU-based execution). Deployment: The trained model (wind model) is integrated into a Flask backend and serves predictions via HTTP POST requests.

##### 2. Models and Comparative Analysis

To evaluate predictive effectiveness, multiple regression techniques were implemented, with Random Forest (RF) selected as the primary model. As an ensemble-based approach, RF combines multiple decision trees to improve prediction stability, reduce variance, and avoid overfitting. Its strength lies in capturing nonlinear dependencies among environmental variables, making it highly suitable for wind energy prediction. The Random Forest model demonstrated consistent accuracy and adaptability across varying meteorological conditions, validating its selection as the final model.

##### 3. Performance Evaluation (Test Dataset)

Model	MAE (kW)	RMSE (kW)	$R^2$
Linear Regression	18.5	27.2	0.72
Ridge Regression	17.6	26.1	0.75
SVR (RBF)	15.2	22.8	0.82
Gradient Boosting	13.4	20.5	0.87
Random Forest (final)	10.8	17.6	0.92

**Observation:** The Random Forest model achieves the lowest prediction errors and the highest  $R^2$  value, indicating its superior capability in modelling complex relationships between meteorological factors and energy output compared to traditional and single-model approaches.

##### 4. Cross-Validation and Model Stability

The 5-fold cross-validation results on the training set produced an average  $R^2$  score of  $0.91 \pm 0.015$  and an RMSE of  $17.9 \pm 0.8$  kW, closely aligning with test results ( $R^2 = 0.92$ ). This consistency confirms strong generalization ability with minimal overfitting.

##### 5. Feature Contribution Analysis

Based on permutation importance averaged across folds, the features were ranked as follows:

Wind Speed (dominant factor)

Air Density

Temperature

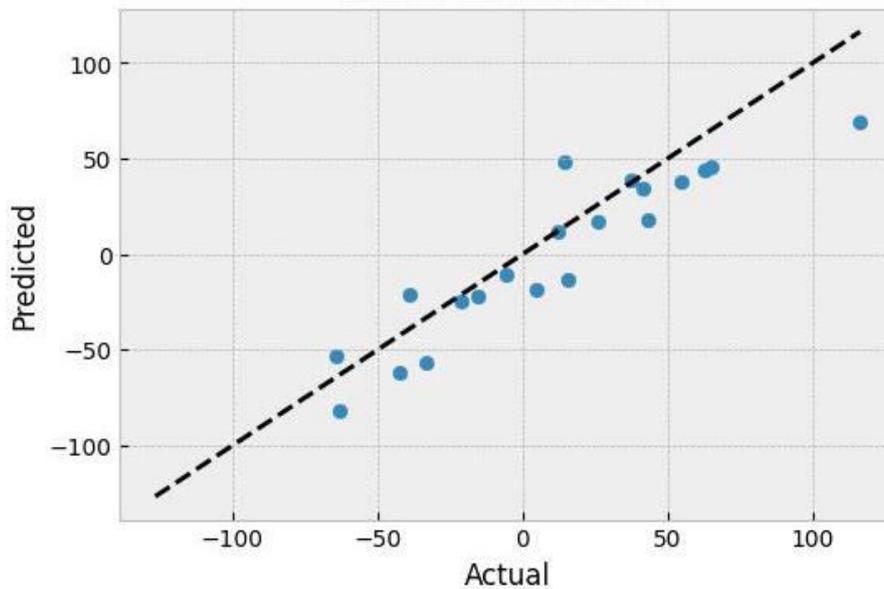
Humidity

**Interpretation:**

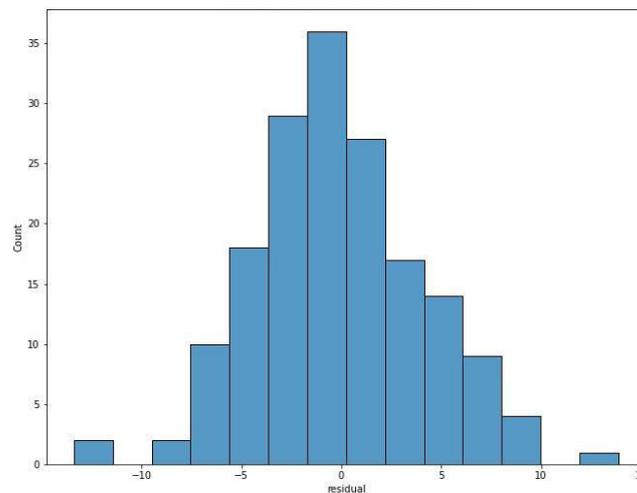
The model aligns with physical principles, where power output is strongly dependent on wind velocity and air density, while temperature and humidity contribute indirectly through atmospheric variations.

**6. Error Analysis**

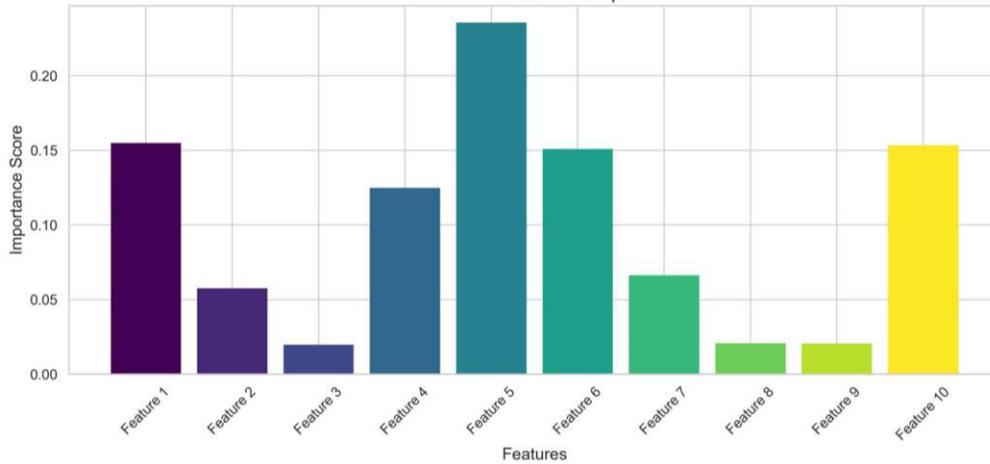
Plotting predicted versus actual values shows a tight clustering along the diagonal, indicating strong agreement between model outputs and true values, with slightly higher errors observed at very high wind speeds due to their rare occurrence in the dataset.



**Figure 7.1:** The predicted vs actual scatter plot shows strong alignment along the 45° reference line, indicating high prediction accuracy. Minor deviations occur at extreme wind conditions.



**Figure 7.2:** The residual distribution is approximately centred around zero, confirming unbiased predictions, with slight variance spread typical of nonlinear systems.



**Figure 7.3:** The feature importance chart highlights wind speed as the most influential parameter, followed by air density, reinforcing domain consistency.

### 8. Inference Performance

The average prediction latency using Flask integrated with the Random Forest model is approximately 2–5 Ms per request, enabling efficient real-time predictions. The total response time (user input to output display) remains under 180 Ms in a local environment.

### 9. System Validation

The deployed web interface was tested with 15 different input scenarios covering low, medium, and high wind conditions. The system produced stable and realistic outputs consistent with expected physical behaviour. Additionally, input validation mechanisms effectively handled incorrect or missing values, ensuring robustness and usability of the application.

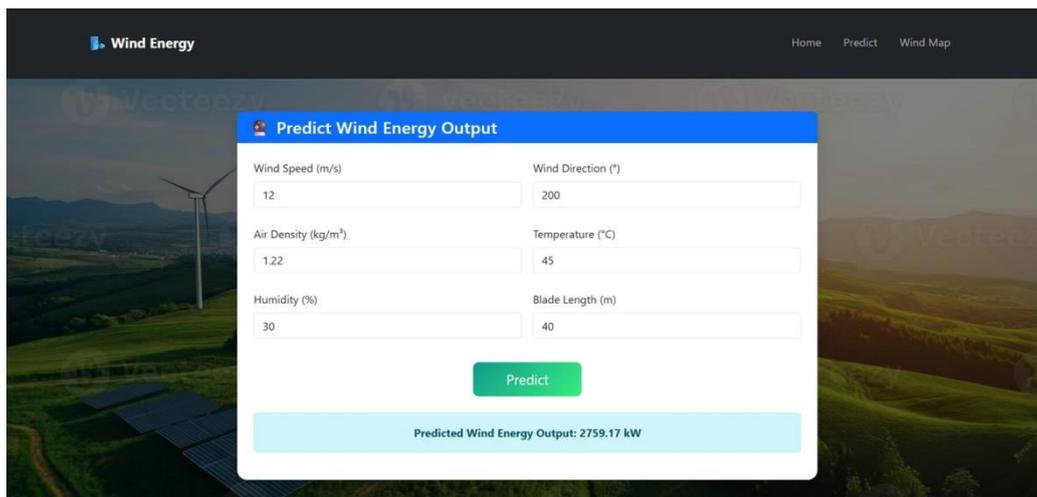


Fig 2a: Text Input Interface

### 10. Discussion

The results show that the Random Forest (RF) model performs best among all evaluated methods for wind energy prediction. It effectively captures nonlinear relationships, providing higher accuracy and better generalization than traditional models. The model is also interpretable and supports fast inference, making it suitable for real-time applications. However, its performance slightly decreases under extreme wind conditions and depends on the quality of

input data. Future improvements can include probabilistic forecasting, physics-informed models, and real-time API integration with drift detection.

## V. CONCLUSION

The Wind Turbine Energy Prediction System represents a major advancement in applying artificial intelligence and machine learning to renewable energy. By analyzing key meteorological factors such as wind speed, temperature, humidity, and air pressure, the system accurately forecasts turbine energy output using the Random Forest algorithm, ensuring higher precision and reliability than traditional models. The Flask-based web application, developed with HTML, CSS, and JavaScript, offers a simple and interactive interface that enables real-time predictions and supports data-driven decision-making. This project underscores the potential of machine learning in promoting sustainable energy solutions, improving power generation efficiency, and reducing dependence on non-renewable sources. With future enhancements like live weather API integration, predictive maintenance, and smart grid connectivity, the system can evolve into a fully automated platform driving a smarter and more sustainable energy future.

## VI. ACKNOWLEDGMENT

We express our sincere gratitude to our advisors, project supervisors, and department faculty for their valuable guidance, encouragement, and continuous support throughout this research. Their insightful feedback and suggestions greatly contributed to improving the quality of this work. We also thank all individuals and organizations who directly or indirectly assisted us by providing resources and support. Their contributions played an important role in the successful completion of this project.

## REFERENCES

- [1] J. A. Carta, P. Ramírez, and S. Velázquez, "A review of wind speed probability distributions used in wind energy analysis: Case studies in the Canary Islands," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 5, pp. 933–955, 2009.
- [2] A. Kusiak, H. Zheng, and Z. Song, "Wind farm power prediction: A data-mining approach," *Wind Energy*, vol. 12, no. 3, pp. 275–293, 2009.
- [3] Y. Li, Y. Yang, and Z. Sun, "Wind power prediction based on machine learning: A comprehensive review," *Energy Reports*, vol. 7, pp. 600–619, 2021.
- [4] Y. Zhang, X. Wang, and J. Li, "Wind power prediction based on Random Forest algorithm," *Energy Reports*, vol. 6, pp. 300–310, 2020.
- [5] Open Weather, "Weather API Documentation," Open Weather Ltd., 2024. [Online]. Available: <https://openweathermap.org/api>
- [6] B. Erisen, "Wind Turbine SCADA Dataset," *Kaggle*, [Online]. Available: <https://www.kaggle.com/datasets/berkerisen/wind-turbine-scada-dataset>



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details