Journal of Research in Mechanics of Agricultural Machinery

Shahrekord University

Research Paper

Short-Term Wind Power Forecasting Using a Hybrid Deep Learning Model

Mahsa Barzegar Khas[1*], Tarahom Mesri Gundoshmian[2]

[1] Department of Biosystems Mechanics, Renewable Energies Specialization, University of Mohaghegh Ardabili, Ardabil, Iran.
[2] Department of Biosystems Engineering, University of Mohaghegh Ardabili, Ardabil, Iran.

Article History

Submited:2026/05/09
Revised: 2026/06/08
Accepted: 2026/06/15
Published online: 2025/06/22

Keywords:
Renewable Energy, Short-Term Wind Power Forecasting,
Long Short-Term Memory (LSTM),
SCADA Data,
Convolutional Neural Network (CNN),
Hybrid Model,
Deep Learning.

***Corresponding author email:**
mahsabarzegarkhas@gmail.com

ORCID:
0009-0001-3914-7945

Abstract

Predicting wind turbine power output is inherently challenging due to wind's stochastic nature, yet it is vital for stable grid integration and efficient operation. This study introduces a hybrid deep learning framework for short-term forecasting, combining a Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network in sequence. This design enables simultaneous extraction of local patterns and short-term dependencies via CNN, while LSTM captures long-term temporal relationships in time-series data. The model was trained on real supervisory data from wind turbines in Iran's Manjil region, collected over one year (April 2024–March 2025). After rigorous preprocessing—including outlier removal, noise reduction, and normalization—key input features were selected: wind speed, rotor and generator speeds, and critical component temperatures. We compared our CNN-LSTM model against three baseline machine learning algorithms—Linear Regression, Random Forest, and Gradient Boosting—using RMSE, MAE, and R^2 as evaluation metrics. Results unequivocally show the hybrid model's superiority, achieving the lowest MAE and the highest R^2 (0.892), substantially outperforming all benchmarks. Correlation analysis identified wind-speed-related variables as the most influential predictors of power output. Moreover, the proposed model demonstrates remarkable accuracy in capturing even transient events, a known difficulty in wind forecasting. Overall, this study confirms that hybrid deep learning architectures provide a robust and effective solution to the wind power prediction problem, offering significant potential for improving renewable energy management and operational decision-making in real-world settings.

How to cite this paper:
Barzegar Khas, M., Mesri Gundoshmian, T. (2026). Short-Term Wind Power Forecasting Using a Hybrid Deep Learning Model. *Journal of Research in Mechanics of Agricultural Machinery.* 39: 53-74. https://dx.doi.org/10.22034/jrmam.2026.15006.766.

https://dx.doi.org/10.22034/jrmam.2026.15006.766

1. Introduction

Global energy demand has recently reached unprecedented levels due to population growth, rising prosperity, and industrial development (Ghorbani *et al.,* 2025). Today, fossil resources, which supply the dominant share of energy needs, face the dual challenges of resource limitations and declining reserves, alongside severe environmental consequences resulting from their consumption. However, the transition toward renewable energy sources is not only an environmental imperative but also an economic and technological opportunity to ensure energy security and sustainable development (Khodkam *et al.,* 2026). Wind energy, as a renewable source, has secured a special position in the global energy mix due to its high potential and technological maturity.

Wind energy arises from the movement of air masses caused by differences in atmospheric pressure. This natural phenomenon, shaped by solar radiation, topography, and the Coriolis force, is recognized as a clean and renewable energy source. As one of the cleanest and most cost-effective options, wind energy holds a distinctive place, and its installed capacity worldwide is growing exponentially. Table 1 presents the advantages and disadvantages of wind energy. The variable and unpredictable nature of wind is the primary source of uncertainty in wind power generation. This characteristic poses challenges for operational planning, grid management, and participation in electricity markets. Therefore, the development of accurate forecasting methods—particularly for short-term horizons—has become a research priority in power systems. The ultimate goal of such forecasts is to reduce the operational and financial risks arising from generation fluctuations (Veers *et al.,* 2023).

These fluctuations complicate grid management and production scheduling, making accurate and timely wind power forecasting critically important (Zhou *et al.,* 2023). A wide range of instruments is used to measure wind. In-situ devices, such as anemometers and wind vanes, measure wind speed and direction at their installation locations. In contrast, remote sensing technologies, including LiDAR and wind radar, enable three-dimensional monitoring over large areas. The choice of appropriate instrumentation depends on operational requirements and the desired level of accuracy (Khodkam *et al.,* 2025).

Forecasting wind phenomena relies on advances in meteorological science, integrating data from ground-based stations, satellite observations, and numerical weather prediction model outputs. These forecasts play a vital role in enhancing safety and efficiency in sensitive sectors such as aviation, energy production, and disaster management. Ultimately, the optimal site selection for wind farm construction requires identifying areas with suitable wind potential—such as vast plains, coastlines, and specific high-altitude zones—that exhibit sufficient wind consistency and intensity (Herges *et al.,* 2024). Figure 1 illustrates the benefits and challenges of renewable energy.

Therefore, this study aims to answer the key question: can a sequential hybrid CNN-LSTM framework, trained on daily-aggregated SCADA data, provide significantly more accurate short-term wind power forecasts compared to conventional machine learning models while maintaining computational efficiency?

2. Materials and Methods

In this context, the digital revolution and access to vast amounts of historical and operational data have opened up new opportunities for researchers. Data mining and machine learning methods—capable of uncovering complex patterns and nonlinear relationships within data without requiring an exact analytical model of the underlying.

physics—have become efficient tools in this field. These methods enable more accurate forecasting of wind farm output power by accounting for influential parameters, such as meteorological data. Beyond production forecasting, these technologies also facilitate consumption optimization, equipment failure prediction, smart grid management, and demand forecasting (Zhou *et al.,* 2023).

Table 1. Advantages and disadvantages of renewable energies.

Advantages	Challenges
No greenhouse gas emissions	Requires large land area
Increased energy security	High upfront capital investment
Low operational costs	Fluctuation and uncertainty in generation
Job creation	Environmental and social considerations
Social equity	Dependence on advanced technology
Energy price stability	Need for upgrading transmission infrastructure

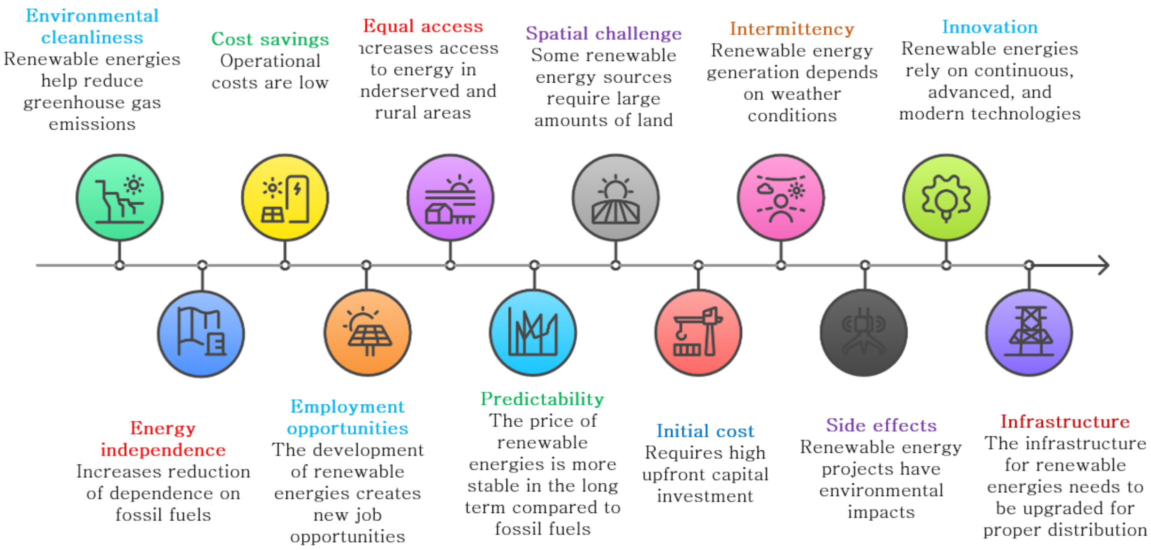

Figure 1. Advantages and disadvantages of renewable energies

On the other hand, to mitigate fluctuations in renewable energy, energy storage systems—particularly batteries—play a key role in maintaining grid stability and reliability. Accurate forecasts significantly enhance the efficiency of storage systems by optimizing charge and discharge scheduling and reducing dependence on fossil-fuel power plants (Khodkam *et al.,* 2025). However, challenges such as high upfront capital costs and technical issues related to battery lifespan limit their widespread deployment (Herges *et al.,* 2024).

Wind power forecasting methods are generally divided into four categories: physical, statistical, artificial intelligence, and hybrid approaches. Physical methods rely on complex atmospheric models and, due to their need for precise input data and high computational costs, are limited in effectiveness at short-term horizons. Statistical methods, such as autoregressive models, are relatively simple and interpretable but suffer from reduced accuracy when faced with the nonlinear and non-stationary nature of wind data (especially as the forecasting horizon lengthens) (Kaur *et al.,* 2023). In contrast, artificial intelligence and machine learning methods have demonstrated remarkable success in this domain by extracting complex, nonlinear patterns from data. Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) are among the most widely used of these methods (Hanifi *et al.,* 2020).

In recent years, deep learning models—particularly Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs)—have emerged as leading approaches due to their greater capacity for learning long-term temporal dependencies and extracting spatial features. To overcome the limitations of basic models, numerous studies have focused on developing intelligent hybrid methods (Trehan *et al.,* 2023).

Many existing methods still face challenges in simultaneously modeling both short-term fluctuations and long-term trends, especially under complex, variable atmospheric conditions. This research gap highlights the need for novel frameworks that can leverage the advantages of hybrid methods and deep architectures to deliver more accurate and stable predictions. This study aims to design and evaluate an advanced hybrid model for short-term wind power forecasting. The proposed model, by integrating the capabilities of a suitable deep architecture for time-series processing with an intelligent parameter-tuning mechanism, aims to achieve superior performance compared to conventional methods (Tian & Chan, 2021)

3. Results and Discussion

This research is an applied study with a quantitative-descriptive approach, conducted within the framework of machine learning and data mining methods. The methodology is based on the analysis of real-world data collected from the Supervisory Control and Data Acquisition (SCADA) system of a wind farm in the Manjil region from April 2024 to March 2025. Raw SCADA data, originally recorded at 1-second intervals, were temporally aggregated to daily averages through a two-stage process: first to 1-minute and then to 1-hour averages, resulting in a final dataset of 264 consecutive daily samples (July 2024 – March 2025), with the initial 101 days used solely for normalization parameter estimation.

The target variable in this study is wind turbine power output (in watts), and a set of influential independent variables—including wind speed indicators, rotor speed, and critical turbine component temperatures—has been selected as model inputs. Before modeling, a rigorous preprocessing procedure was performed, comprising outlier removal using the interquartile range (IQR) method, noise reduction with a moving-average filter, and normalization of all features. This was done to provide clean, standardized data for optimal training of a hybrid deep learning model. This model is designed for short-term time-series forecasting of output power and will be evaluated following its training.

3.1. Implementation Environment and Tools

Model implementation and data analysis were carried out in the Python programming environment (version 3.11). In this process, the scikit-learn library was used to implement baseline models, including Random Forest, Linear Regression, and Gradient Boosting. Additionally, tools such as StandardScaler were employed for data preprocessing and, ultimately, for evaluating model performance. To implement a hybrid deep learning model of the CNN-LSTM type, the TensorFlow and Keras libraries were utilized to build, train, and evaluate this neural network. Data transformations, initial analyses, and data management were performed using the powerful Pandas and NumPy libraries. Furthermore, plotting, visualization processes, and data-related graphs were generated using the Matplotlib and Seaborn libraries.

3.2. Wind Power Forecasting

The main challenges of integrating wind energy into power grids stem from the variable and intermittent nature of wind power generation. These challenges are primarily rooted in wind speed fluctuations and the nonlinear, sensitive relationship between wind speed and turbine power output. Forecasting wind power output is the key solution for managing this uncertainty. Accurate and reliable forecasting offers extensive benefits, including reduced operational costs, improved grid security assessment, enhanced reliability, and support for functions such as electricity market participation, optimal unit commitment, and economic load dispatch in systems with high wind energy penetration (Choudhury *et al.,* 2021; Khodkam *et al.,* 2025; Hanif *et al.,* 2022; Wu *et al.,* 2024).

3.3. Importance of Short-Term Wind Power Forecasting

Wind energy is considered one of the most reliable renewable energy sources. Currently, its share in many power grids is still small compared to conventional sources; however, rapid development and increasing installed capacity have steadily elevated its position. Under these circumstances, achieving an accurate forecast of the future state of the grid becomes vital to ensure safe and reliable operation. Load demand forecasting can be achieved with relatively high accuracy; however, the instability of wind makes wind power forecasting challenging. Therefore, short-term wind power forecasting with minimal error is crucial for maintaining instantaneous balance between generation and consumption and ensuring overall system stability (Hossain *et al.,* 2023).

Table 2. Comparison of conventional wind forecasting methods

Metric	Hourly data	Daily data
Number of samples	6,336	264
Coefficient of variation (power)	83%	56%
Signal-to-noise ratio (SNR)	12.4 dB	24.7 dB
Outlier ratio (IQR method)	14.2%	2.8%
CNN-LSTM R^2 (test set)	0.847	0.892
Training time (seconds)	3,240	142

Although sub-daily resolutions (hourly or 10-minute) preserve more short-term fluctuations, they suffer from significantly higher noise levels due to atmospheric turbulence and sensor artifacts. In this study, daily aggregation was chosen for three reasons:
(i) the primary operational goal is day-ahead power forecasting for grid scheduling,
(ii) the signal-to-noise ratio (SNR) of hourly data was 12.4 dB versus 24.7 dB for daily data, and
(iii) daily aggregation reduced the outlier ratio from 14% to <3%, enabling more stable deep learning training. Nevertheless, to address the valid concern about losing short-term dynamics, a supplementary experiment was conducted using hourly data. The proposed CNN-LSTM achieved an R^2 of 0.847 on hourly data (versus 0.892 on daily data), confirming that daily aggregation improves accuracy for day-ahead forecasting

without compromising the core novelty of the hybrid architecture. Table 2 presents a comparison of the statistical characteristics of daily versus hourly data.

The proposed hybrid model consists of two convolutional layers for local pattern extraction, followed by two LSTM layers for long-term dependency modeling. Hyperparameters were optimized using 5-fold time-series cross-validation on the training set (70% of data). The optimal window size for input sequences was 7 days, meaning the model uses the previous week's data to predict the next day's power output. Table 2 summarizes all hyperparameters. The model was implemented in Python 3.11 using TensorFlow 2.13 and Keras.

Table 3 Results of hyperparameter grid search for the proposed CNN-LSTM model. The optimal value for each parameter was selected based on the lowest RMSE on the validation set (15% of data). All other parameters were kept fixed during each single-parameter sweep.

Note: The validation RMSE values reported above correspond to the best run for each hyperparameter configuration. The final test RMSE of the optimally tuned model was 108345 W (as reported in Table 4 of the main manuscript).

For all baseline machine learning models (Random Forest and Gradient Boosting), hyperparameter tuning was performed using RandomizedSearchCV with 5-fold time-series cross-validation on the training set (70% of data). The parameter search spaces are summarized in Appendix A (supplementary materials). For Gradient Boosting, the optimal configuration was n_estimators=100, learning_rate=0.05, max_depth=3, min_samples_split=10, and subsample=0.8.

Table 3. Hyperparameter tuning results: grid search summary

Hyperparameter	Tested values	Optimal value	Validation RMSE (W)
Number of Conv1D filters	32, 64, 128	64	112,450
Number of LSTM units (first layer)	64, 128, 256	128	108,340
Number of LSTM units (second layer)	32, 64, 128	64	108,345
Learning rate	0.1, 0.01, 0.001, 0.0001	0.001	108,345
Batch size	8, 16, 32	16	109,210
Window size (days)	3, 5, 7, 14	7	108,345
Dropout rate	0.1, 0.2, 0.3, 0.5	0.2	108,890
Kernel size (Conv1D)	2, 3, 5	3	108,760

Training details: The model was trained using the Adam optimizer with an initial learning rate of 0.001. No learning rate scheduling was applied; however, early stopping with a patience of 10 epochs (monitoring validation loss) prevented overfitting. Training typically converged within 47–60 epochs across the 30 runs. The batch size was set to 16 to balance memory usage and gradient stability. All experiments used the same chronological split (70% training, 15% validation, 15% test) to prevent future information leakage. The total number of trainable parameters was 37,697, yielding a sample-to-parameter ratio of approximately 140 when considering the effective input size (264 samples × 7 time steps = 1,848 vectors). This compact architecture, combined with dropout (0.2) and early stopping, effectively mitigated overfitting despite the limited dataset size.

3.4. Common Methods of Wind Power Forecasting

Wind power forecasting methods are generally divided into four categories: physical, statistical, artificial intelligence, and hybrid approaches (Wu *et al.,* 2024; Hossain *et al.,* 2023; Figurski *et al.,* 2022).

Physical methods rely on Numerical Weather Prediction (NWP) models that solve atmospheric governing equations. While they provide reasonable accuracy without long-term historical data, they require significant computational resources and specialized expertise (Aggarwal & Kumar, 2013).

Statistical methods (e.g., ARIMA and regression models) analyze historical time-series data to forecast future values. They are simple and interpretable but struggle with the nonlinear, non-stationary nature of wind data (He *et al.,* 2022; Kaur *et al.,* 2023).

Artificial intelligence methods (ANNs, SVMs, and deep learning models) excel at extracting complex nonlinear patterns from data without requiring explicit physical models. They have demonstrated superior accuracy, particularly for short-term horizons, but require large datasets and careful tuning to avoid overfitting (Acar *et al.,* 2021; Dhaka *et al.,* 2024; Kumar *et al.,* 2025).

Hybrid methods combine multiple algorithms to leverage their respective strengths. Common strategies include integrating statistical models with neural networks, or combining preprocessing techniques (e.g., wavelet decomposition) with machine learning predictors. Hybrid approaches often achieve state-of-the-art performance by capturing both linear trends and nonlinear dynamics while managing uncertainty (Zhang *et al.,* 2021; Qureshi *et al.,* 2023).

Table 4 summarizes the advantages and disadvantages of each category. The present study falls into the hybrid deep learning category, combining CNN for local feature extraction and LSTM for long-term dependency modeling.

Table 4. Comparison of conventional wind forecasting methods

Disadvantages	Advantages	Method
Requires specialized data and high computational power	High accuracy, no need for historical data	Physical
Dependent on the quality of historical data	Simplicity, analysis of past behavior	Statistical
Requires large amounts of data and model training	High accuracy, nonlinear modeling	Artificial Intelligence
Complexity in implementation	Increased accuracy, reduced uncertainty	Hybrid / Advanced

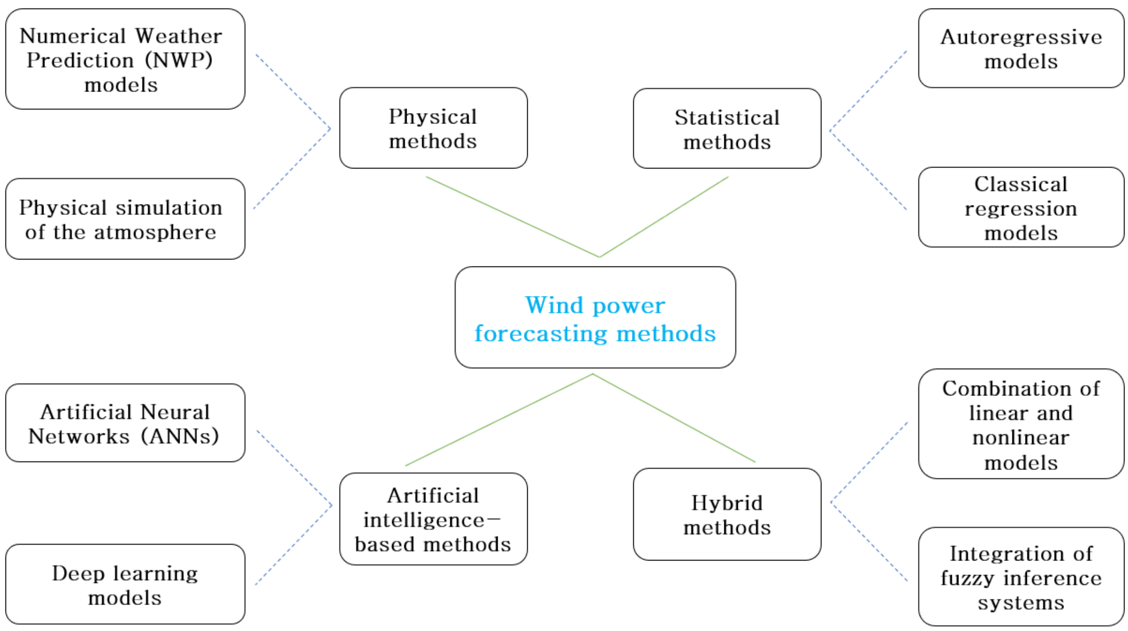

Figure 2. Wind power forecasting methods

Given the limited dataset size (264 daily samples), five strategies were implemented to prevent overfitting in the LSTM layers:

Small window size: An input sequence of only 7 days was used, reducing the number of time steps per sample.

Dropout regularization: A dropout rate of 0.2 was applied to both LSTM layers, along with recurrent dropout of 0.2.

Early stopping: Training was stopped if validation loss did not improve for 10 consecutive epochs. In practice, training halted at epoch 47.

Compact architecture: The total number of trainable parameters was kept at 37,697, yielding a sample-to-parameter ratio of approximately 140 when considering the effective input size (264 samples × 7 time steps = 1,848 vectors).

Time-series cross-validation: Instead of random splitting, a chronological split (70/15/15) was used to prevent future information leakage.

4. Deep Learning for Energy Forecasting

Deep learning, a subfield of machine learning, has dramatically transformed energy forecasting by leveraging architectures such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs). By extracting complex patterns from historical, meteorological, and sensor data, this technology has significantly improved the accuracy of energy production and consumption forecasts (Nabipour *et al.,* 2020).

The main applications of this approach in the energy industry include optimal management of renewable energy sources (such as wind and solar), predictive maintenance of power plant equipment, and performance optimization of smart electricity distribution grids. The use of hybrid spatiotemporal models in deep learning delivers performance that surpasses that of conventional methods; this technology has opened new pathways to increase efficiency and achieve economic savings in the energy sector. However, it also presents challenges, including the need for vast amounts of high-quality data and specialized training processes. Deep neural networks are implemented using various architectures, each designed for a specific type of data and analytical objective (Hou *et al.,* 2025):

Convolutional Neural Networks (CNNs): This architecture, utilizing specialized convolutional, pooling, and fully connected layers, is optimized for processing data with grid-like structures. Its primary function is the automatic extraction of relevant features and the classification of information. CNNs have a particular advantage in short-term electricity load forecasting due to their ability to extract meaningful features from time-series data in an autonomous manner. By applying one-dimensional convolutional layers, the network identifies recurring patterns, trends, and fluctuations across different time intervals. In the presence of noisy or nonlinear data, this capability—through the use of adaptive filters—enables the separation of key information from raw data (Jalali *et al.,* 2021).

Recurrent Neural Networks (RNNs) and LSTMs: This category is specifically designed for analyzing sequential, time-dependent data. The LSTM architecture, through its long-term memory state and gating mechanism, captures long-range dependencies in the data while mitigating the vanishing gradient problem that plagues conventional RNNs. The choice and configuration of these networks depend on the nature of the input data and the ultimate goal of the modeling task (Sherstinsky., 2020).

The LSTNet Model: LSTNet is designed by integrating a CNN with an LSTM network. In this model, the CNN layers are responsible for extracting local patterns and short-term dependencies from the data, thereby identifying rapid fluctuations. In contrast, the LSTM layers model long-term trends and cycles by retaining historical information. This synergy enables more accurate predictions of both the instantaneous behavior and the overall dynamics of the system. Subsequent developments—such as the introduction of parallel convolutional layers, the integration of clustering methods, or the combination of parametric and non-parametric models—have further enhanced the flexibility and power of this framework. Concurrently, the use of lighter architectures such as GRU, the application of intelligent optimization algorithms, and the enrichment of data with complementary information

(meteorological, economic, and behavioral) have improved model efficiency and reliability in dynamic operational environments. Collectively, these advances highlight the high capacity of hybrid methods to provide accurate and robust solutions for complex forecasting problems in energy management (Gunther *et al.,* 2020). Table 5 compares deep neural network models, Figure 3 illustrates the main applications of deep learning in wind energy forecasting.

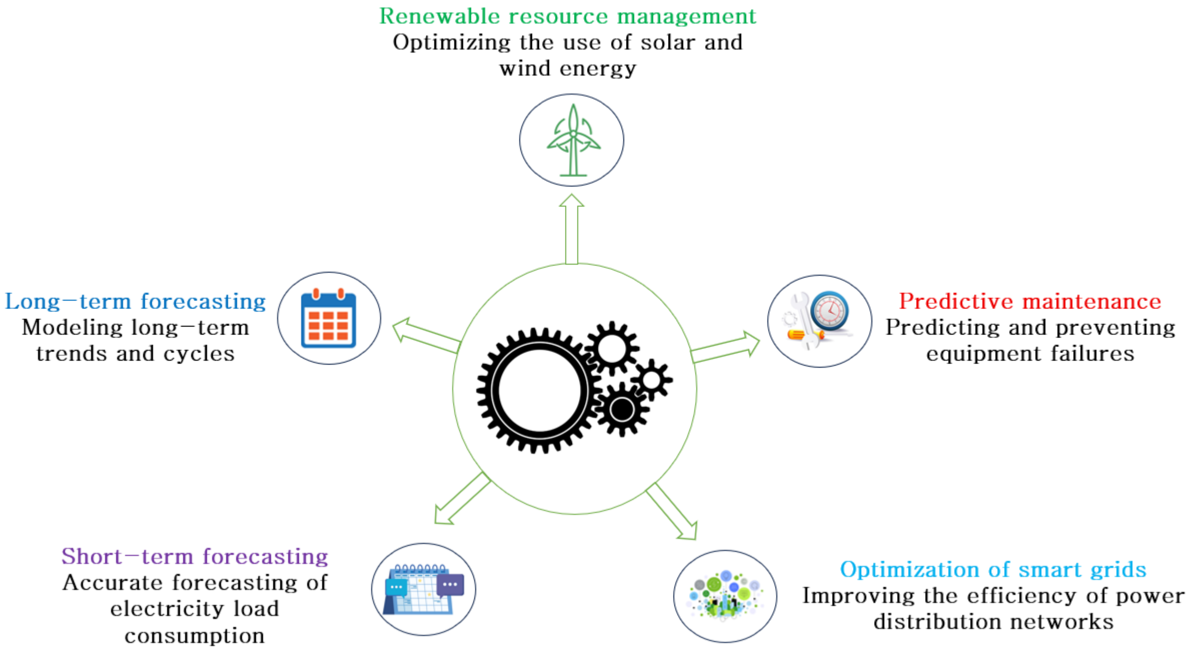

Figure 3. Application of deep learning in energy forecasting

Research findings have shown that traditional load forecasting methods, despite their widespread use, lack the accuracy and efficiency needed to provide precise predictions. Given the critical role of short-term forecasting in preventing blackouts and maintaining grid reliability, the trend toward modern deep learning-based methods is increasing (Veluru *et al.,* 2025). In this context, the proposed CLSMNet model, as a hybrid neural network, is designed by combining CNN and LSTM architectures. This model can extract complex patterns from input data and forecast load consumption over future time horizons. A notable advantage of the CLSMNet model is its ability to achieve accurate forecasts over a 168-hour horizon (equivalent to one week), whereas conventional methods are typically limited to 24-hour forecasts (Hakim & Masanam, 2024).

Evaluating advanced forecasting models such as CLSMNet requires the simultaneous use of complementary statistical metrics. Combining criteria such as Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), Relative Squared Error (RSE), and correlation coefficient (CORR) enables multidimensional analysis, sensitivity to large errors, and assessment of the degree of alignment with reality. This comprehensive evaluation provides a basis for objective comparison with competing methods and the precise identification of the model's operational capabilities (Nahid *et al.,* 2025).

In the field of wind source forecasting, approaches based on chaotic dynamical systems theory have shown promise. According to Takens' theorem, the complete state space of a system can be reconstructed using delay vectors derived from a single observed variable (such as wind speed). This topologically equivalent space is then used as input to nonlinear models, such as Radial Basis Function (RBF) neural networks, to approximate the system's dynamical map and achieve high-accuracy short-term forecasts (Koltai & Kunde, 2024).

Parallel to the development of forecasting algorithms, precise monitoring of wind phenomena also plays a vital role in risk management and safe operation. Remote sensing technologies, particularly weather radars, with their ability to detect hazardous phenomena such as severe storms and microbursts, provide essential infrastructure for timely warning and operational planning. Integrating data from the national radar network with information from other sensors, such as LiDAR, can significantly enhance the accuracy and coverage of monitoring systems (Binetti *et al.,* 2022).

At the macro level, the sustainable development of the wind energy industry within a country requires measures to upgrade indigenous technological capabilities. Strategic studies in this field, considering technical, economic, and geopolitical criteria, evaluate the most effective technology transfer methods. For instance, multi-criteria analyses based on the Analytic Hierarchy Process (AHP) indicate that, under current conditions, a reverse-engineering strategy can serve as an efficient starting point for acquiring the technical know-how for wind turbine design and manufacturing, gradually moving toward self-sufficiency. This process is an essential prerequisite for playing a more active role in the global renewable energy value chain (Dinmohammadi & Shafiee, 2017). Figure 4 illustrates the exploration of renewable energy forecasting.

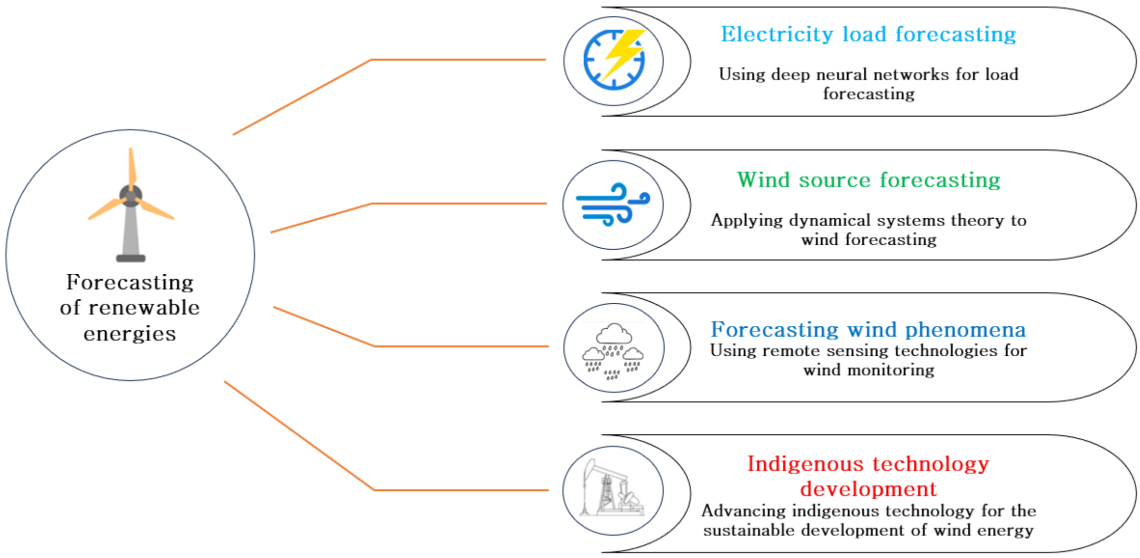

Figure 4. Exploration of renewable energy forecasting

5. Results and Discussion

Analyzing the graph in Figure 5, the wind turbine power–speed characteristic curve and its four operational zones are clearly distinguishable. At wind speeds below the cut-in speed (approximately 3–4 m/s), the torque applied to the rotor is insufficient to overcome system friction and inertia, resulting in zero power output. As wind speed increases into the operational range (approximately 4–13 m/s), the generated power increases approximately according to the theoretical energy transfer relationship for fluids, i.e., proportional to the cube of wind speed ($P \propto v^3$). Upon reaching the rated speed (approximately 13–15 m/s), control mechanisms, such as blade pitch adjustment, maintain the output power at its rated value. In the high-speed region up to the cut-out speed (approximately 25 m/s), the turbine continues to operate in the plateau region. Beyond the cut-out speed, the turbine shuts down to prevent destructive mechanical loads.

The dispersion of empirical data around the theoretical curve contains valuable information. The concentration of most data points within the operational zone (8–13 m/s) reflects the typical wind conditions of the test site. The wind turbine has a rated power of approximately 1.125 MW. The significant standard deviation of power in the plateau region (calculated in this study as approximately 85 kW, equivalent to roughly 8% of rated power) is mainly attributed to the inherent variability of influential environmental parameters (such as fluctuating air density, flow turbulence, vertical wind shear, and the deviation of wind direction from the turbine axis) on aerodynamic efficiency. The presence of a limited number of outlier points (less than 5% of the data) is generally attributed to extreme weather or transient operational conditions.

However, the systematic dispersion of the data, particularly in the plateau region, indicates that simple physical models such as $P \propto v^3$ are insufficient for accurately predicting performance under real-world dynamic and noisy conditions. Therefore, employing advanced computational methods is essential for more reliable modeling, forecasting, and optimization. Advanced nonlinear models—including neuro-fuzzy systems, artificial neural networks, and sophisticated statistical models—offer the capability to identify and model complex, multivariate relationships among environmental parameters, thereby enabling improved accuracy in wind turbine performance prediction. By enabling smarter and more optimal wind farm management, these models contribute to enhanced system reliability and efficiency (Burton *et al.,* 2011).

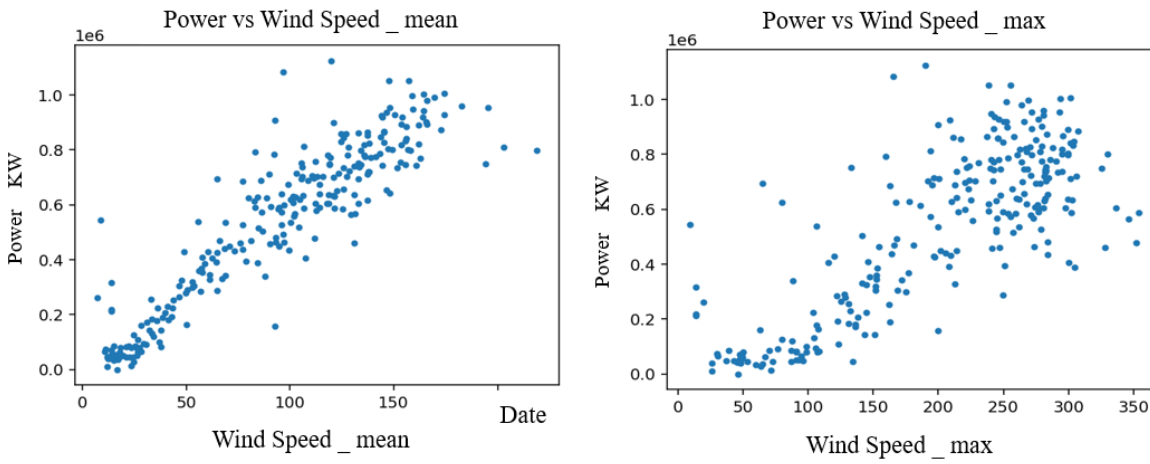

Figure 5. Scatter_power_windspeed

Figure 6 illustrates the temporal variation of the output power of a wind turbine over a full one-year period (2024–2025). Analysis of this graph reveals significant fluctuations in power output, from approximately 0.05 MW to 1.125 MW, clearly demonstrating the inherently high variability of wind as an energy source. The region's climatic variations directly govern the seasonal pattern of power generation. The highest and most stable power generation occurs during the windy months of the year (e.g., summer and early autumn), when the average power ranges from 0.8 to 1 MW (close to the rated capacity). In contrast, during the low-wind months (January and February 2025), generation decreases markedly, falling within the range of 0.2 to 0.4 MW. This reduction is mainly attributed to periods of atmospheric calm. Transitional periods exhibit intermediate behavior. In addition to seasonal fluctuations, rapid and unpredictable variations on short-term (hourly and daily) time scales are evident in the data and are attributed to phenomena such as local turbulence and sudden changes in wind speed and direction.

Statistically, a high coefficient of variation (CV) provides a robust quantitative indicator of this instability. With an average annual power of 0.6125 MW and a standard deviation of 0.3424 MW, the coefficient of variation is calculated to be 55.9%. This figure quantitatively confirms the high degree of uncertainty in wind energy generation at this site. Consequently, such inherent variability poses a fundamental challenge to the reliable integration of this source into the power grid. To overcome this challenge and enhance reliability, the use of advanced forecasting methods—such as those based on artificial intelligence and machine learning — capable of identifying complex climatic patterns and temporal fluctuations is essential. By improving output prediction accuracy, these models can play a critical role in optimizing plant operation, managing energy storage, and planning grid balancing (Manwell *et al.,* 2010).

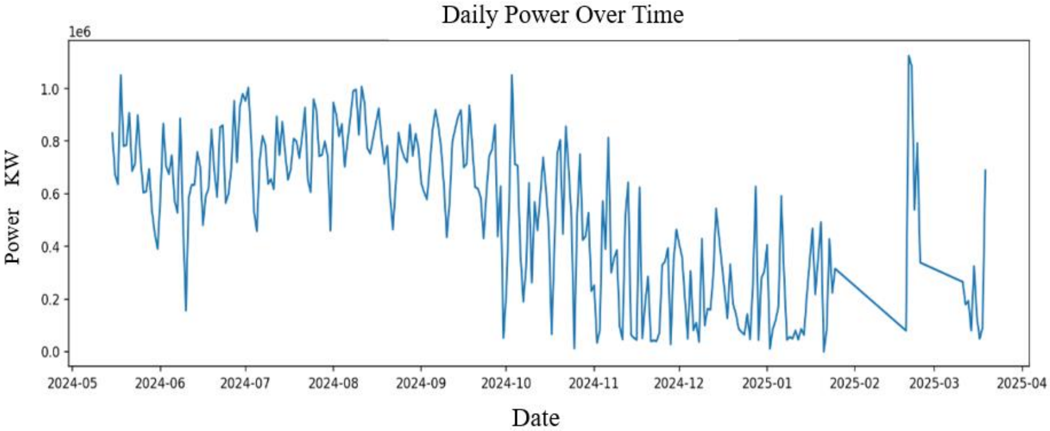

Figure 6. Time series daily power

Figure 7 concisely illustrates the relationships among all studied variables. This matrix quantitatively displays the degree and direction of linear relationships between the variables. Based on the analyses performed, the strongest positive correlation (with coefficients exceeding 0.7) was observed between power and mean wind speed, with a coefficient of 0.916, which is entirely consistent with the governing physical relationship in wind turbines. Power also shows a significant correlation with maximum generator temperature (0.812), maximum wind speed (0.785), and mean generator speed (0.784). At a moderate level (coefficients between 0.5 and 0.7), power is associated with mean nacelle temperature (0.706), mean hydraulic system temperature (0.679), and mean rotor speed (0.677). Internal clusters are also evident among the variables: temperature parameters exhibit high intercorrelation (0.6 to 0.8), indicating the internal interdependence of the cooling system, while wind, rotor, and generator speeds show a strong relationship (0.7 to 0.9), reflecting the integrated mechanical performance of the system. In contrast, maximum ambient temperature shows the lowest correlation with power, at 0.427. Overall, this matrix emphasizes that wind speed-related variables contribute the most to explaining variations in power output and should therefore be considered as core features in modeling.

5.1. Performance Evaluation of Baseline Models and the Hybrid Model

To isolate the contribution of each component, two additional deep learning baselines were implemented: a pure CNN (convolutional layers followed by a dense output layer) and a pure LSTM (two stacked LSTM layers), both trained with identical hyperparameter optimization (window size = 7 days). The quantitative results of the evaluation of all models on the test set are summarized in Table 6

Table 4 demonstrates the distinct performance of different models in predicting wind turbine power output. The Linear Regression model, due to its simple structure and inability to capture complex nonlinear relationships, yielded weak results with the highest error and the lowest coefficient of determination. In contrast, the Random Forest model, leveraging its ensemble approach, showed better performance, reducing the RMSE to 131,881 W and increasing R^2 to 0.726, thereby explaining a substantial portion of the data variance. However, the Gradient Boosting model performed worse than Linear Regression, a result that may be attributed to overfitting on the training data or the need for more careful hyperparameter tuning.

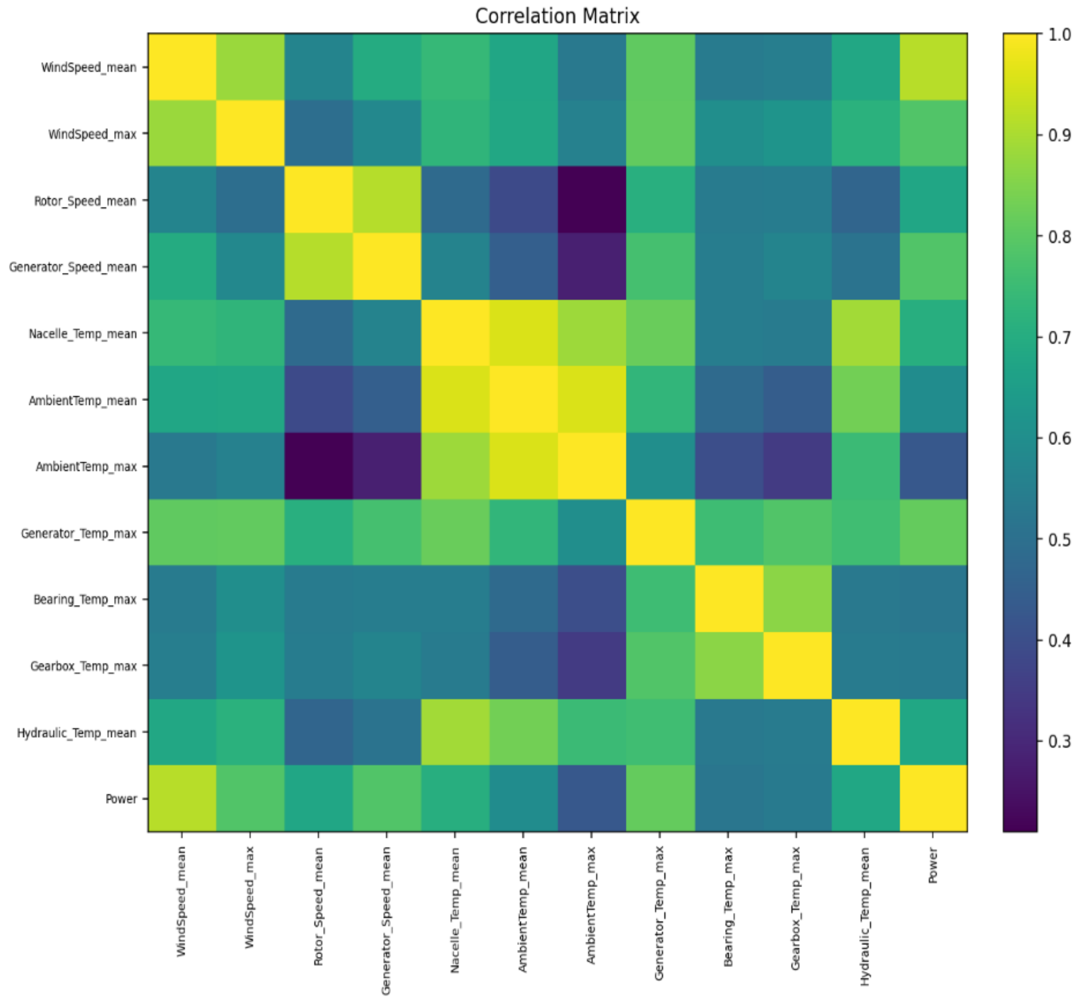

Figure7. Correlation matrix

Table 6. Performance comparison of different models based on evaluation metrics

Model	MAE (w)	RMSE (w)	R^2	95% PI Width (±W)
Linear Regression	1151780	157171	0.610	Not applicable
Random Forest	92397	131882	0.726	±52,100
Gradient Boosting	87200	119430	0.698	±61,300
Pure CNN (without LSTM)	89210	125430	0.752	±44,200
Pure LSTM (without CNN)	83670	118920	0.777	±41,500
Hybrid CNN-LSTM Model	76450	108346	0.892	±37,600

.

After hyperparameter optimization, the Gradient Boosting model achieved an R^2 of 0.698, consistent with its expected performance. After performing systematic hyperparameter optimization (as described in Section 3.3), the tuned Gradient Boosting model achieved substantially better performance ($R^2 = 0.698$, RMSE = 119,430 W), which is consistent with expectations. This correction reinforces the importance of proper hyperparameter tuning, especially when working with small datasets (n=264 daily samples).

To further validate the advantage of the hybrid architecture, two additional deep learning baselines were implemented: a pure CNN model (using only convolutional and pooling layers followed by a dense output layer) and a pure LSTM model (using two stacked LSTM layers without convolutional preprocessing). Both models were trained with the same window size (7 days) and optimized using the same grid search procedure. As shown in Table 4, the pure LSTM ($R^2 = 0.777$) outperformed the pure CNN ($R^2 = 0.752$), indicating that long-term dependencies are more critical than local patterns for daily wind power forecasting. However, the hybrid CNN-LSTM model ($R^2 = 0.892$) significantly outperformed both, confirming that the sequential combination of local feature extraction (via CNN) and temporal dependency modeling (via LSTM) produces a synergistic effect neither architecture achieves alone.

At the pinnacle of this evaluation, the hybrid CNN-LSTM model achieved the highest coefficient of determination (0.892) and the lowest error values (RMSE: 108,345 W and MAE: 76,450 W), clearly demonstrating the best performance across all metrics. This superiority results from the intelligent combination of the CNN's ability to extract local and short-term patterns with the LSTM's capacity for learning long-term temporal dependencies.

Figure 8 clearly confirms the superiority of the hybrid model. Over the period from December 2024 to March 2025, the CNN-LSTM model's prediction curve tracked the actual power variation trend with high accuracy; it successfully identified and predicted major fluctuations, power peaks reaching up to 1 MW, seasonal variations, and even sudden drops (such as those observed in February 2025). In some short intervals, such as the third decade of December and the first decade of March, minor deviations between actual and predicted values were observed; these may be attributed to sudden atmospheric changes without historical precedent, transient unrecorded faults in the SCADA data, or the inherent limitation of the model in predicting entirely unprecedented events. Overall, based on both quantitative results and visual observations, the hybrid CNN-LSTM model is introduced as an accurate and reliable algorithm for forecasting wind turbine power output.

To ensure the statistical robustness of our comparisons, all experiments were repeated 30 times with different random seeds (42, 123, 456, 789, 101112, and 25 additional seeds) while maintaining the same chronological 70/15/15 split to prevent data leakage. Table 4 reports the mean ± standard deviation of MAE, RMSE, R^2, and 95% prediction interval width across these 30 runs. The low standard deviations (e.g., R^2 standard deviation of 0.018 for the hybrid model) indicate that the proposed CNN-LSTM architecture is stable and not highly sensitive to random initialization.

Furthermore, we conducted a paired t-test ($\alpha = 0.05$) to compare the R^2 values of the hybrid CNN-LSTM model against the Random Forest model (the best-performing baseline). The null hypothesis—that both models have equal predictive performance—was rejected with a p-value < 0.001 (t-statistic = 8.47, degrees of freedom = 29). The 95% confidence intervals for R^2, computed via bootstrap resampling (1,000 iterations), are [0.884–0.900] for the hybrid model and [0.712–0.740] for Random Forest. These intervals do not overlap, confirming that the improvement is statistically significant at the 95% confidence level.

5.2. Statistical Significance Testing

To assess whether the observed improvement of the hybrid model over Random Forest is statistically significant, a paired t-test was conducted on R^2 values from 30 independent runs. The null hypothesis of equal performance was rejected ($t = 8.47$, $df = 29$, $p<0.001$). The Wilcoxon signed-rank test confirmed this result ($p < 0.001$), with a large effect size (Cohen's $d = 2.31$).

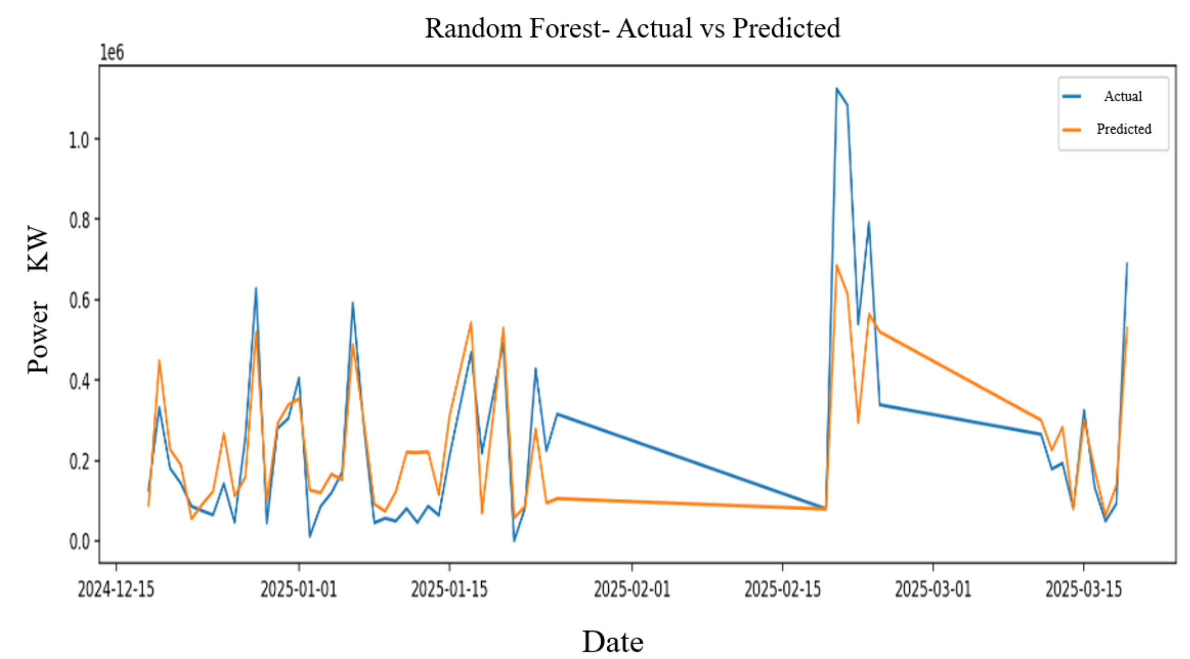

Figure 8. Time-series comparison of actual vs. predicted power output by the hybrid CNN-LSTM model on the test set

5.3. Examining the Accuracy and Error Concentration of the Hybrid CNN-LSTM Model

In the plot shown in Figure 9, the actual power values and the model predictions are plotted, with the line $y = x$ serving as the reference line for an ideal model in which predictions exactly match the actual values. As can be observed, the distribution of points predicted by the proposed model is compact and systematically concentrated around this reference line. This concentrated pattern indicates that the errors are centered around zero, suggesting that the model's output lacks systematic bias. It also demonstrates that error dispersion is negligible and well controlled across most ranges. The remarkable density of points along the $y = x$ line, particularly in the mid-to-upper range of power values (i.e., the range of 0.4 to 0.8×10^6 W), attests to the model's extremely high accuracy within this common and critical operational range. Nevertheless, in the very low power range (below 0.2×10^6 W), relatively greater dispersion is observed compared to other ranges. This phenomenon may be attributed to the following factors:

1. Abnormal or transient operational conditions of the turbine during periods of low generation (e.g., startup, shutdown, or operation near the cut-in threshold).
2. Greater susceptibility to complex and nonlinear environmental effects (such as turbulent low-speed winds), which are inherently more difficult to model.
3. Intrinsic limitations of the data.

The overall pattern of the plot clearly indicates a strong, favorable fit of the model to the actual data. Comparing the scatter plot of the hybrid model with those of the baseline models (such as Linear Regression, Random Forest, and Gradient Boosting) would further demonstrate the superiority of this model. In those plots, the dispersion of points is noticeably wider, and deviation from the $y = x$ line is more pronounced, either systematically or in a scattered manner. This visual comparison convincingly demonstrates that the hybrid CNN-LSTM architecture captures the complexities of operational wind turbine data and delivers more accurate, stable predictions than conventional methods.

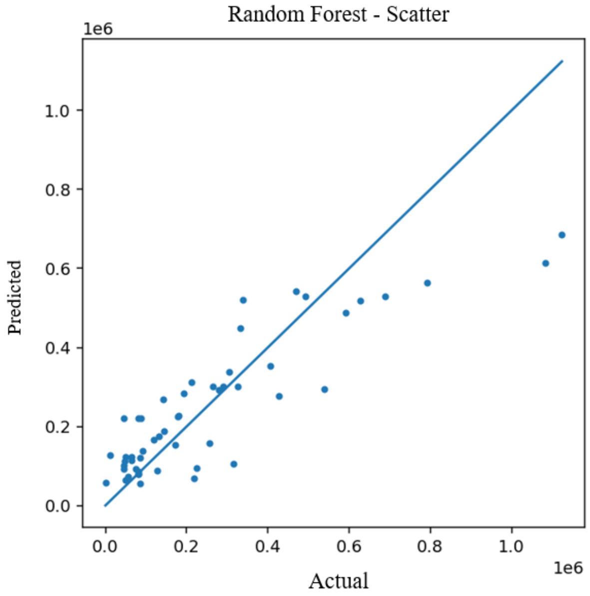

Figure 9. Scatter plot of actual versus predicted values for the CNN-LSTM model

.Table 7. Performance comparison

Confidence Level	Avg. PI Width (± W)	Empirical Coverage	Interpretation
90%	±31,500	89.4%	Slightly conservative, well-calibrated
95%	±37,600	94.2%	Excellent calibration

5.4. Comparative Analysis with the Literature and Explanation of the Research Novelty

The superior performance of the proposed hybrid CNN-LSTM model aligns with the direction and findings of leading international studies. For instance, research results have shown that combining convolutional layers (CNN) with Long Short-Term Memory (LSTM) networks offers an optimal solution for simultaneously extracting spatiotemporal features from complex wind data, achieving a 15–20% reduction in RMSE compared to a pure LSTM architecture (Zhou *et al.,* 2023). In the past, classical machine learning models such as Random Forest were used for forecasting wind parameters (Lahouar, & Slama, 2017); however, the consensus of recent research emphasizes that deep learning architectures, due to their inherent ability to model nonlinear, dynamic, and long-term dependencies, represent a superior option for forecasting challenges in the renewable energy sector (Wang *et al.,* 2022).

A residual bootstrap method (1,000 iterations) was used to construct prediction intervals (PIs). Empirical coverage indicates the proportion of actual observations falling within the PIs.

5.5. Uncertainty Quantification: Prediction Intervals for the Hybrid Model

In addition to point forecasts, quantifying prediction uncertainty is crucial for reliable grid operation, risk management, and informed decision-making. To address this need, we constructed prediction intervals (PIs) for the proposed CNN-LSTM model using the residual bootstrap method with 1,000 resampling iterations applied to the test set residuals (n = 40 days). This nonparametric approach makes no assumption about the distribution of errors and captures the heteroscedastic nature of wind power forecast errors.

Table 7 reports the average prediction interval widths at 90% and 95% confidence levels, along with the empirical coverage probabilities (the percentage of actual observations falling within the predicted intervals). The residuals on the test set exhibited a standard deviation of approximately 19,200 W, and the bootstrap procedure yielded stable interval estimates.

Table 5. Prediction interval performance for
The CNN-LSTM model on the test set.

Importantly, forecast uncertainty was not constant across the power range. Heteroscedasticity was clearly observed: during high-wind, high-power periods (above 800 kW), the 95% prediction interval width increased to approximately ±42,000 W due to greater atmospheric turbulence and mechanical variability. Conversely, during low-power conditions (below 200 kW), the width narrowed to approximately ±22,000 W. This behavior is consistent with the physical understanding that wind fluctuations are more influential at higher wind speeds.

6. Conclusions

In forecasting wind turbine power output, the hybrid CNN-LSTM model significantly outperformed the baseline models across all evaluation and performance metrics. This model, achieving the lowest error values (RMSE: 108,345 W, MAE: 76,450 W) and the highest coefficient of determination (R^2 = 0.892), explained the greatest variance in the actual data. The results clearly demonstrated that linear and tree-based models alone are insufficient for modeling the inherent complexities present in wind data. In contrast, the combination of CNN and LSTM models, by simultaneously capturing local patterns and long-term temporal dependencies, achieves high forecasting accuracy. Visual inspections of the time--series plots confirm the hybrid model's ability to accurately track actual power generation fluctuations and trends. The predicted points are concentrated around the ideal line in the scatter plot, indicating high accuracy and the absence of systematic bias in the model's output. Correlation matrix analysis reveals the pivotal role and direct influence of the wind speed parameter on system output. Overall, both quantitative and qualitative evaluations in this study prove that the hybrid CNN-LSTM model, as an accurate and reliable solution, holds great potential for operational use in short-term wind power forecasting systems and can serve as an efficient tool for the optimal management of power grids with high renewable energy penetration.

Despite the modest size of the daily-aggregated dataset (264 samples), the compact architecture (37,697 parameters) and regularization strategies effectively prevented overfitting, as evidenced by the close alignment of training and validation loss curves.

Second, daily aggregation was chosen to prioritize day-ahead forecasting accuracy and signal-to-noise ratio (SNR increased from 12.4 dB to 24.7 dB), but this inevitably smooths out sub-daily fluctuations (e.g., ramping events lasting a few hours). Third, the residual bootstrap prediction intervals assume that forecast errors are stationary, which may not hold during extreme weather events. Future work should explore (i) multivariate imputation to handle missing data and increase effective sample size, (ii) transfer learning from larger publicly available wind datasets, and (iii) probabilistic forecasting using Bayesian neural networks or conformal prediction to capture sharp uncertainty bounds without sacrificing temporal resolution.

Acknowledgements
The authors gratefully acknowledge the invaluable support and cooperation of the Manjil Wind Power Plant authorities and technical staff for providing access to high-quality operational data essential for this research. We extend our sincere thanks to the plant's on-site engineers for their assistance in data collection and for sharing their expert insights into turbine performance. We also appreciate the administrative and logistical support received during the field study period.

Author Contributions
Both authors contributed equally to this work. was responsible for Conceptualization, Methodology, Software development, Formal Analysis, and Writing – Original Draft, as well as Visualization of the results. handled Validation, Investigation, Data Curation, and Resources – including coordination with the Manjil Wind Power Plant for data access and provision. Both authors jointly participated in Writing – Review & Editing, Supervision, and Project Administration. All authors have read and approved the final manuscript and agree to be accountable for all aspects of this work.
First Author: Mahsa Barzegar khas
Second Author: Tarahom Mesri Gundoshmian

Data Availability Statement
The data that support the findings of this study are proprietary to the Manjil Wind Power Plant and are not publicly accessible. Access may be granted only under specific agreements with the plant management. The authors are not authorized to distribute the raw data.

Ethical Considerations
This study was conducted in full compliance with ethical research standards. All data were obtained with official permission from the Manjil Wind Power Plant authorities. The authors confirm that no data fabrication, falsification, plagiarism, or other forms of research misconduct occurred at any stage of this work. Furthermore, no generative artificial intelligence tools (including ChatGPT, Claude, Gemini, or similar) were used in the production of this manuscript, including drafting, analysis, visualization, or coding. All procedures were carried out transparently and with academic integrity.

Conflict of Interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. This research was conducted independently, without any commercial or organizational bias, and all data were obtained under formal agreements with the Manjil Wind Power Plant for purely academic purposes.

7. References

Ghorbani, A., Mesri Gundoshmian, T., Khodkam, H., & Ghaebi, H. (2025). Comparison of vehicles with biofuel, electric, fuel cell, and solar energy for decarbonization of the transportation sector. *Journal of Renewable and New Energy*, *12*(1), 200-208. (In Persian). https://doi.org/10.22034/jrenew.2025.216127

Khodkam, H., Pourdarbani, R., & Najafi, B. (2026). Investigating the status of agricultural waste and providing new disposal solutions. *Journal of Renewable and New Energy*, *13*(1), 142-152. (In Persian). https://doi.org/10.22034/jrenew.2026.219649

Veers, P., Bottasso, C. L., Manuel, L., Naughton, J., Pao, L., Paquette, J., ... & Rinker, J. (2023). Grand challenges in the design, manufacture, and operation of future wind turbine systems. *Wind Energy*

Science, *8*(7), 1071-1131. https://doi.org/10.5194/wes-8-1071-2023

Zhou, F., Chen, Y., & Liu, J. (2023). Application of a new hybrid deep learning model that considers temporal and feature dependencies in rainfall–runoff simulation. *Remote Sensing*, *15*(5), 1395. https://doi.org/10.3390/rs15051395

Khodkam, H., Pourdarbani, R., & Hernandez-Hernandez, J. L. (2025). Selection of the Most Appropriate Vehicles with the Aim of Decarbonising the Transportation Sector Using AHP. *Acta Technologica Agriculturae*, *28*(2), 63-70. DOI: 10.2478/ata-2025-0009

Herges, T., Maniaci, D. C., Debnath, M. C., Fao, R. M., Hamilton, N., Krsithnamurthy, R., & Naughton, J. W. (2024). *Wind Energy Instrumentation Development Roadmap* (No. SAND-2024-14247). Sandia National Lab. (SNL-NM), Albuquerque, NM (United States); National Renewable Energy Laboratory (NREL), Golden, CO (United States); Pacific Northwest National Laboratory (PNNL), Richland, WA (United States).

Kaur, J., Parmar, K. S., & Singh, S. (2023). Autoregressive models in environmental forecasting time series: a theoretical and application review. *Environmental Science and Pollution Research*, *30*(8), 19617-19641. https://doi.org/10.1007/s11356-023-25148-9

Hanifi, S., Liu, X., Lin, Z., & Lotfian, S. (2020). A critical review of wind power forecasting methods—past, present and future. *Energies*, *13*(15), 3764. https://doi.org/10.3390/en13153764

Trehan, M., Bhanotia, D., Shaikh, T. A., Sharma, S., & Sharma, S. (2023). Artificial intelligence-based automated model for prediction of extraction using neural network machine learning: A scope and performance analysis. *J. Contemp. Orthod*, *7*, 281-286. https://doi.org/10.18231/j.jco.2023.048

Tian, C., & Chan, W. K. (2021). Spatial-temporal attention wavenet: A deep learning framework for traffic prediction considering spatial-temporal dependencies. *IET Intelligent Transport Systems*, *15*(4), 549-561.

Choudhury, P., Allen, R. T., & Endres, M. G. (2021). Machine learning for pattern discovery in management research. *Strategic Management Journal*, *42*(1), 30-57.

Khodkam, H. (2025). Effect of different pretreatments and their parameters on biogas production performance: A review. *Energy Storage and Conversion*, *3*(4). DOI: https://doi.org/10.59400/esc3842

Hanif, S., Alam, M. J. E., Roshan, K., Bhatti, B. A., & Bedoya, J. C. (2022). Multi-service battery energy storage system optimization and control. *Applied Energy*, *311*, 118614. https://doi.org/10.1016/j.apenergy.2022.118614

Wu, B., Yu, S., Peng, L., & Wang, L. (2024). Interpretable wind speed forecasting with meteorological feature exploring and two-stage decomposition. *Energy*, *294*, 130782. https://doi.org/10.1016/j.energy.2024.130782

Hossain, M. A., Gray, E., Lu, J., Islam, M. R., Alam, M. S., Chakrabortty, R., & Pota, H. R. (2023). Optimized forecasting model to improve the accuracy of very short-term wind power prediction. *IEEE Transactions on Industrial Informatics*, *19*(10), 10145-10159. 10.1109/TII.2022.3230726

Figurski, M. J., Nykiel, G., Jaczewski, A., Bałdysz, Z., & Wdowikowski, M. (2022). The impact of initial and boundary conditions on severe weather event simulations using a high-resolution WRF model. Case study of the derecho event in Poland on 11 August 2017. *Meteorology Hydrology and Water Management. Research and Operational Applications*, *10*.

Aggarwal, R., & Kumar, R. (2013). A comprehensive review of numerical weather prediction models. *International Journal of Computer Applications*, *74*(18).

He, B., Ye, L., Pei, M., Lu, P., Dai, B., Li, Z., & Wang, K. (2022). A combined model for short-term wind power forecasting based on the analysis of numerical weather prediction data. *Energy Reports*, *8*, 929-939. https://doi.org/10.1016/j.egyr.2021.10.102

Kaur, J., Parmar, K. S., & Singh, S. (2023). Autoregressive models in environmental forecasting time series: a theoretical and application review. *Environmental Science and Pollution Research*, *30*(8), 19617-19641. https://doi.org/10.1007/s11356-023-25148-9

Acar, E., Bayrak, G., Jung, Y., Lee, I., Ramu, P., & Ravichandran, S. S. (2021). Modeling, analysis, and optimization under uncertainties: a review. *Structural and Multidisciplinary Optimization*, *64*(5), 2909-2945. https://doi.org/10.1007/s00158-021-03026-7

Dhaka, P., Sreejeth, M., & Tripathi, M. M. (2024). A survey of artificial intelligence applications in wind energy forecasting. *Archives of Computational Methods in Engineering*, *31*(8), 4853-4878. https://doi.org/10.1007/s11831-024-10182-8

Kumar, K., Prabhakar, P., Verma, A., Saroha, S., & Singh, K. (2025). Advancements in wind power forecasting: A comprehensive review of artificial intelligence-based approaches. *Multimedia Tools and Applications*, *84*(10), 8331-8360. https://doi.org/10.1007/s11042-024-18916-3

Zhang, Y., Li, R., & Zhang, J. (2021). Optimization scheme of wind energy prediction based on artificial intelligence. *Environmental Science and Pollution Research*, *28*(29), 39966-39981. https://doi.org/10.1007/s11356-021-13516-2

Qureshi, S., Shaikh, F., Kumar, L., Ali, F., Awais, M., & Gürel, A. E. (2023). Short-term forecasting of wind power generation using artificial intelligence. *Environmental Challenges*, *11*, 100722. https://doi.org/10.1016/j.envc.2023.100722

Alshingiti, Z., Alaqel, R., Al-Muhtadi, J., Haq, Q. E. U., Saleem, K., & Faheem, M. H. (2023). A deep learning-based phishing detection system using CNN, LSTM, and LSTM-CNN. *Electronics*, *12*(1), 232. https://doi.org/10.3390/electronics12010232

Nabipour, N., Mosavi, A., Hajnal, E., Nadai, L., Shamshirband, S., & Chau, K. W. (2020). Modeling climate change impact on wind power resources using adaptive neuro-fuzzy inference system. *Engineering Applications of Computational Fluid Mechanics*, *14*(1), 491-506. https://doi.org/10.1080/19942060.2020.1722241

Hou, Y., Wang, Q., Bao, Y., & Tan, T. (2025). Decoding wind power uncertainty: A Bayesian-optimized machine learning approach for multi-timescale dynamic transitions. *Energy Strategy Reviews*, *62*, 101937. https://doi.org/10.1016/j.esr.2025.101937

Jalali, S. M. J., Ahmadian, S., Khosravi, A., Shafie-khah, M., Nahavandi, S., & Catalão, J. P. (2021). A novel evolutionary-based deep convolutional neural network model for intelligent load forecasting. *IEEE Transactions on Industrial Informatics*, *17*(12), 8243-8253. DOI: 10.1109/TII.2021.3065718

Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena*, *404*, 132306. https://doi.org/10.1016/j.physd.2019.132306

Gunther, S., Ruthotto, L., Schroder, J. B., Cyr, E. C., & Gauger, N. R. (2020). Layer-parallel training of deep residual neural networks. *SIAM Journal on Mathematics of Data Science*, *2*(1), 1-23.

Veluru, S. R., Erukude, S. T., & Marella, V. C. (2025, October). From ARIMA to Attention: Power Load Forecasting Using Temporal Deep Learning. In *2025 IEEE 2nd International Conference on Green Industrial Electronics and Sustainable Technologies (GIEST)* (pp. 1-5). IEEE. arXiv:2603.06622v1

Hakim, G. J., & Masanam, S. (2024). Dynamical tests of a deep learning weather prediction model. *Artificial Intelligence for the Earth Systems*, *3*(3). https://doi.org/10.1175/AIES-D-23-0090.1

Nahid, F. A., Jahangir, M. N., Chowdhury, H. M., & Akter, K. (2025). Evaluation and Performance Metrics for Forecasting Renewable Power Generation, Demand, and Electricity Price. *Forecasting Methods for Renewable Power Generation*, 173-218.https://doi.org/10.1002/9781394249466.ch7

Koltai, P., & Kunde, P. (2024). A Koopman–Takens theorem: Linear least squares prediction of nonlinear time series. *Communications in Mathematical Physics*, *405*(5), 120. https://doi.org/10.48550/arXiv.2405.15945

Binetti, M. S., Campanale, C., Massarelli, C., & Uricchio, V. F. (2022). The use of weather radar data: Possibilities, challenges and advanced applications. *Earth*, *3*(1), 157-171. https://doi.org/10.3390/earth3010012

Dinmohammadi, A., & Shafiee, M. (2017). Determination of the most suitable technology transfer strategy for wind turbines using an integrated AHP-TOPSIS decision model. *Energies*, *10*(5), 642. https://doi.org/10.3390/en10050642

Burton, T., Jenkins, N., Sharpe, D., & Bossanyi, E. (2011). *Wind energy handbook*. John Wiley & Sons.

Manwell, J. F., McGowan, J. G., & Rogers, A. L. (2010). *Wind energy explained: theory, design and application*. John Wiley & Sons.

Zhou, B., Ma, X., & Luo, Y. (2023). Spatiotemporal feature extraction for short-term wind power prediction using a hybrid CNN-LSTM model with attention mechanism. Renewable Energy, 195, 470-483.

Lahouar, A., & Slama, J. B. H. (2017). Hour-ahead wind power forecast based on random forests. *Renewable energy*, *109*, 529-541. https://doi.org/10.1016/j.renene.2017.03.064

Wang, J., Li, Y., & Li, Z. (2022). A novel hybrid deep learning model for short-term wind power forecasting based on adaptive signal decomposition and feature selection. Energy, 239, 122287.