Sound & Vibration 2025, 59(3), 2059.
https://doi.org/10.59400/sv2059

Article

Impact of vibration on wind turbine efficiency and LSTM-based power conversion prediction

Aarif Alutaybi, Chafaa Hamrouni*

Department of Computer Sciences, Taif University-Khurma University College, Khurma 2935, Kingdom of Saudi Arabia
*** Corresponding author:** Chafaa Hamrouni, cmhamrouni@tu.edu.tn

CITATION

Alutaybi A, Hamrouni C. Impact of vibration on wind turbine efficiency and LSTM-based power conversion prediction. Sound & Vibration. 2025; 59(3): 2059.
https://doi.org/10.59400/sv2059

ARTICLE INFO

Received: 15 November 2024
Accepted: 4 June 2025
Available online: 29 June 2025

Abstract: During the long-term operation of wind power generation systems, the impact of mechanical vibration on energy conversion efficiency is often overlooked. Existing studies mostly use vibration signals as a means of fault warning, lacking a systematic analysis of the quantitative relationship between vibration characteristics and power generation efficiency. This study, based on Supervisory Control and Data Acquisition (SCADA) data and high-frequency vibration monitoring from a wind farm, extracts vibration features from both time domain (e.g., root mean square, peak value, skewness, and kurtosis) and frequency domain (e.g., dominant frequency and spectral energy ratio). Through Pearson and Spearman correlation analyses, as well as a comparative time series analysis of high-vibration intervals (revealing an average efficiency drop of 3.5%), it is demonstrated that intensified vibrations significantly reduce generation efficiency and increase output fluctuations. Furthermore, a dual-layer LSTM prediction model is proposed, integrating wind speed, wind direction, temperature, and vibration features. The training process is optimized using a sliding window strategy, Dropout regularization, and early stopping. On the test set, the model achieves an RMSE of 0.035 and a MAPE of 3.6%, outperforming support vector machines (SVM), random forests, and single-layer GRU models by 20%–40% in accuracy. Finally, an integrated "monitoring–prediction–warning–control" framework is proposed to support real-time deployment and intelligent operation and maintenance (O&M), offering a practical solution for wind farm health management and O&M optimization.

Keywords: wind turbine vibration; power generation efficiency; LSTM; SCADA; prediction model

1. Introduction

With the continuous optimization of the global energy structure and the in-depth promotion of the "dual carbon" strategy, wind energy is increasingly attracting widespread attention [1]. Wind power generation has been rapidly promoted around the world due to its low carbon emissions and resource sustainability [2]. However, wind turbines will inevitably produce various forms of structural vibrations when they operate under complex and changeable wind conditions and environmental conditions for a long time [3]. These vibrations may originate from factors such as blade aerodynamic load fluctuations, main shaft imbalance, gear meshing errors, etc., which not only affect the structural safety and operating stability of the unit, but also have potential interference with energy conversion efficiency [4]. Therefore, further research on the mechanism of vibration impact on wind power generation system operation performance is important in improving the utilization rate of wind energy and guaranteeing the reliable operation of the equipment.

Existing research mainly focuses on the impact of climate-related factors on wind energy conversion devices, and has made some progress in wind energy conversion efficiency analysis and wind turbine fault diagnosis [5,6]. However, systematic quantitative studies on the dynamic coupling between vibration behavior and power generation efficiency are still lacking, especially in the context of actual wind farm operation data, there is still an obvious gap on how to identify the mechanism of vibration's effect on energy conversion efficiency. In particular, in the context of actual wind farm operation data, there is still a clear gap in how to identify the mechanism of vibration on energy conversion efficiency. On the other hand, Supervisory Control and Data Acquisition (SCADA) systems are widely used to provide massive monitoring data for wind power operations, promoting data-driven high-precision modeling and performance prediction [7]. On this basis, the development of deep learning technology has provided a new methodological path for building a wind energy conversion efficiency prediction model.

Based on the above context, this study aims to impact of wind turbine vibration characteristics on power generation efficiency, and constructs a model based on the Long Short-Term Memory (LSTM) network to achieve high-precision efficiency prediction. By analyzing key variables from real SCADA data, including nacelle vibration, wind speed, wind direction, and output power, this work extracts time-domain and frequency-domain vibration features that are relevant to efficiency. These features are then used to quantify the relationship between vibration behavior and power output. In addition, by leveraging the modeling strengths of deep learning, the study achieves dynamic prediction of energy conversion efficiency, thereby providing both theoretical insight and practical guidance for condition monitoring, efficiency assessment, and optimized maintenance in wind power systems.

This research holds significance on both theoretical and practical fronts. Theoretically, it fills a gap in the literature by examining the influence of vibration parameters on wind energy conversion efficiency and contributes to the broader paradigm of multiphysical field coupling analysis in wind power systems. Practically, the LSTM-based prediction model offers algorithmic support for intelligent monitoring and proactive intervention in wind farms, thus demonstrating clear engineering applicability. Moreover, the findings are expected to support turbine health monitoring, early warning of abnormal vibrations, and the formulation of cost-effective operation strategies for wind farms.

To achieve the research objectives, this thesis systematically explores the core issues of wind farm operation. First, professional data cleaning is performed on the original data collected by the SCADA system to remove outliers and fill in missing values. At the same time, key vibration characteristics and environmental variables are extracted in combination with the equipment operation characteristics. Then, statistical methods are used for in-depth analysis to quantify the degree of correlation between vibration characteristics and power generation efficiency. Then, a conversion efficiency prediction model is developed based on the characteristics of the LSTM network, and the performance is compared with traditional algorithms such as support vector machines. The model output is then analyzed in combination with actual working conditions to explore the efficiency change law under different vibration conditions. Finally, based on the actual scenario of the wind farm, a practical strategy covering real-time prediction and operation and maintenance optimization is proposed to provide support for improving operational efficiency.

The thesis is structured into six sections: Section 1 introduces the research background, objectives, significance, and content; Section 2 reviews relevant studies on turbine vibration analysis, efficiency modeling, and deep learning prediction methods; Section 3 details the dataset, preprocessing, feature extraction methods, and LSTM model architecture; Section 4 presents data analysis and experimental results, evaluating the performance of various models; Section 5 offers a detailed discussion of the results, including physical interpretations, model limitations, and practical implications; and Section 6 concludes the study and outlines directions for future research.

2. Related work

2.1. Factors influencing wind turbine performance

The power generation efficiency of wind turbines is affected by the coupling of multiple factors such as wind energy resource characteristics, aerodynamic design, mechanical state degradation, and environmental and operation and maintenance [8–11]. Among them, wind speed, wind direction, air density and meteorological conditions are the core environmental operating parameters that affect wind turbine efficiency [12]. For example, Peter et al. found that weather conditions (especially wind speed) can significantly affect the reliability of modern wind turbines, and different wind speeds have different effects on the reliability of different turbine subcomponents. Due to more severe weather conditions, the reliability problem of offshore wind power facilities is more prominent than that of onshore wind turbines [13].

In addition, seasonal fluctuations in air density also have a direct impact on wind energy capture efficiency [14]. Under high temperature and high humidity environmental conditions, air density decreases, resulting in a decrease in wind turbine output power; in cold seasons, the blade surface may increase aerodynamic resistance due to ice or frost, thereby significantly reducing the efficiency of energy conversion.

In addition to environmental factors, the impact of the mechanical structure of the wind turbine on its performance cannot be ignored [15]. Studies have found that abnormal changes in blade or shaft vibration frequency are often highly correlated with a decrease in power generation efficiency. Frequent high-amplitude vibrations can also aggravate material fatigue damage, shorten equipment life, and increase maintenance costs [16].

Therefore, combining vibration monitoring with efficiency evaluation not only helps to reveal the inherent mechanism between mechanical dynamic behavior and energy conversion efficiency, but also provides a theoretical basis and technical support for implementing predictive maintenance and optimizing wind turbine operation strategies.

2.2. Wind turbine data monitoring and health diagnosis

Wind turbine vibration monitoring is an important means to achieve equipment health management and status assessment [17]. With the expansion of wind farm scale and the extension of operating life, traditional manual inspection and regular maintenance alone can no longer meet the dual requirements of equipment reliability and economy. At present, wind turbines are generally equipped with SCADA systems to collect and integrate various parameter information during wind turbine operation in real time, including vibration amplitude, temperature, current, voltage, speed and power generation [18]. Through the fusion analysis of these multi-source heterogeneous data, the visual display and trend analysis of wind turbine operation status can be effectively realized, providing basic support for equipment status identification and anomaly detection.

In recent years, researchers have proposed a method based on the Normal Behavior Model (NBM). By establishing a typical parameter change law model in the early stage of wind turbine operation or the healthy stage, abnormal patterns that deviate from the normal trajectory can be detected in real time on this basis, thereby realizing early warning of faults and identification of performance degradation trends [19,20].

To improve the diagnostic accuracy and system intelligence, machine learning and deep learning methods are widely used in vibration signal feature extraction and fault identification tasks [21,22]. Ogaili et al. [23] proposed a fusion of the ReliefF algorithm and the KNN classifier to significantly improve the accuracy of wind turbine blade crack detection; Zhao et al. [24] constructed an adaptive threshold model based on the theory of deep autocoder and extreme value to realize the early fault identification of SCADA data.

In summary, vibration monitoring and health diagnosis are evolving towards intelligence, integrating multi-source data to build prediction models, which helps improve the stability and power generation efficiency of wind power systems.

2.3. Research status of wind turbine power generation efficiency prediction

Wind turbine power generation efficiency prediction methods have evolved from statistical models to machine learning models [25]. Early models (ARMA & ARIMA) had limited effectiveness due to their difficulty in handling the nonlinear volatility of wind speed. Subsequently, support vector machines (SVM) combined with SCADA multi-source data improved prediction accuracy and adaptability [26,27].

Recently, deep learning models have shown remarkable strengths in wind turbine power prediction. Studies have shown that combining LSTM with machine learning can effectively improve prediction accuracy and reduce error indicators such as RMSE and MAE [28]. At the same time, LSTM variants with self-attention mechanisms (such as LSTM + Attention) have more advantages in identifying complex time-dependent patterns, and have become a hot topic in current wind energy prediction research [29].

Although a large number of studies have been conducted on the application of vibration signals in wind turbine fault diagnosis, the research on directly using vibration data for power generation efficiency prediction modeling is still relatively limited, and few studies have jointly modeled vibration characteristics with actual wind turbine power output. Therefore, it is of great significance to deeply explore the impact of vibration on wind energy conversion efficiency.

2.4. Research gaps and innovations

Although current wind power research has made significant progress in fault diagnosis and performance prediction, there are still several research gaps. Most literature uses vibration signals as a means of early fault warning, but lacks in-depth discussion of the quantitative relationship between vibration signals and the actual power generation efficiency of wind turbines. In particular, research that integrates SCADA operational data with high-frequency vibration signals to construct a unified modeling framework for elucidating the direct impact of vibration on energy conversion efficiency is still limited.

In addition, existing power prediction models mostly focus on environmental or electrical factors such as wind speed and rotation speed, and have not fully considered the role of dynamic behaviors such as mechanical vibration in efficiency changes, which limits the integrity and accuracy of the prediction model.

Therefore, this study proposes an LSTM time series modeling method that integrates vibration and environmental parameters to quantitatively reveal the impact mechanism of vibration characteristics on power generation efficiency. By combining sliding windows with frequency domain/time domain feature engineering, the model structure and hyperparameter configuration are optimized to construct a high-performance model with actual prediction capabilities. The study will be empirically verified based on multiple wind turbine SCADA and vibration monitoring data to test the generalization ability and engineering application potential of the proposed method, filling the gap in the existing model in modeling the relationship between mechanical vibration and power output.

3. Methodology

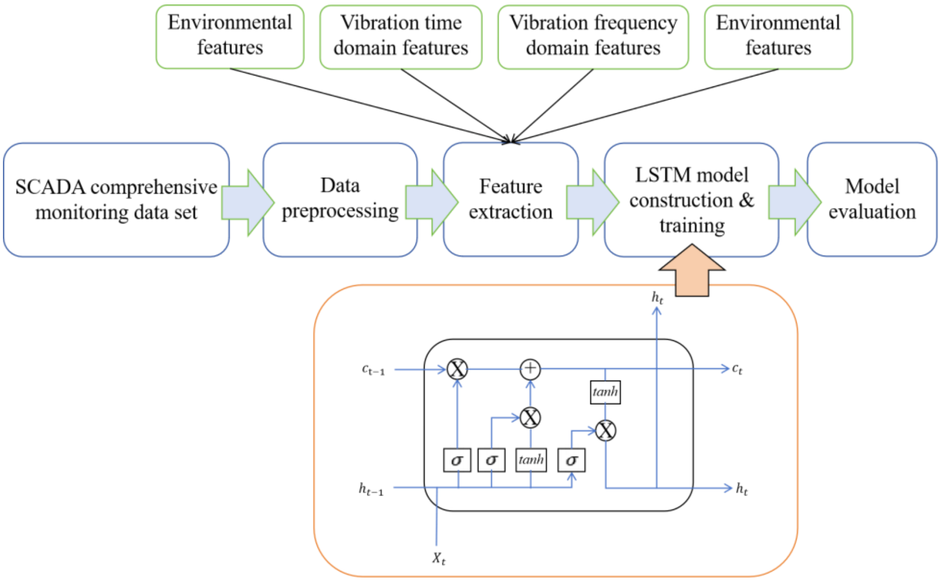

Figure 1. Overall flow chart of the experiment.

This section will introduce in detail the data set source and preprocessing process, feature engineering method, LSTM model construction process and model evaluation index system used in this study. The overall process aims to build a deep learning model that integrates vibration and environmental characteristics to achieve accurate prediction of wind turbine power generation efficiency and explore the impact of vibration characteristics on its performance. The overall experimental process is shown in **Figure 1**.

3.1. Dataset selection and preprocessing

This study uses the wind turbine SCADA comprehensive monitoring dataset included in the PMC data platform (2023) [30]. This dataset covers a set of long-term operation information of a typical onshore wind turbine, including key variables such as mechanical vibration acceleration, three-dimensional wind speed, wind direction, air temperature, and electrical output power, with a time resolution of 10 min. Given the study's focus on the coupling relationship between vibration characteristics and energy conversion efficiency, the data selection phase prioritized variables including acceleration response signals from the blades and main shaft, wind speed and direction, and power output as the main subjects of analysis.

During the preprocessing stage, missing values were imputed using linear interpolation combined with a moving average technique. Outliers were identified and removed using boxplot analysis and physical threshold constraints. To enhance the stability of model training, all input variables were normalized using the Z-score standardization method, ensuring that all features were scaled to a common numerical range for effective feature learning.

3.2. Feature engineering

To more effectively extract efficiency-related information from vibration signals, this study adopts a combined approach of time-domain and frequency-domain feature extraction. In the time domain, key features such as root mean square (RMS), peak value, skewness, and kurtosis of vibration acceleration are extracted. In the frequency domain, features including dominant frequency, spectral energy distribution, and band energy ratio are derived using Fast Fourier Transform (FFT) to capture potential anomalies in structural responses. These vibration features are computed over sliding time windows to construct multi-scale feature sequences.

In terms of environmental factors, in addition to the basic wind speed, wind direction and temperature, combined features such as wind speed change rate and interaction terms between wind speed and vibration (such as wind speed × RMS) are also constructed to enhance the model's ability to characterize nonlinear interactive relationships.

3.3. LSTM prediction model construction

This study uses the LSTM as the core structure of the prediction modeling to capture the temporal dependency between wind turbine vibration, environmental changes and power output. The model structure consists of an input layer, one or two layers of LSTM hidden units, and a fully connected output layer to achieve the modeling of the nonlinear mapping between multidimensional time series features and target variables [31].

The input data is a multidimensional sequence in a sliding window of length T, that is,

$$X = \{x_1, x_2, \ldots, x_T\}, x_t \in R^d \tag{1}$$

where d represents the feature dimension, T is the time step, and the model output is the predicted value of the target variable (such as power output) at the current moment $\hat{y}_T$.

The core mechanism of the LSTM unit includes the following update formula:

Forget gate:

$$f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f) \tag{2}$$

(1) Input gate and candidate state:

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\tilde{C}_t = \tanh(W_C \times [h_{t-1}, x_t] + b_C) \tag{4}$$

(2) Unit state and output gate:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t, \tag{5}$$

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o), \tag{6}$$

$$h_t = o_t \times \tanh(C_t) \tag{7}$$

The final output is mapped to the predicted value through a fully connected layer:

$$\hat{y}_T = W_{out} \times h_T + b_{out} \tag{8}$$

In terms of hyperparameter settings, the initial time step is set to 10, the number of hidden units is 64 or 128, the optimizer is Adam, the learning rate is set to 0.001, and Dropout regularization is used to prevent overfitting. During the training process, the sliding window method is used to amplify the time series samples, combined with 5-fold cross-validation to improve the generalization ability, and the early stopping mechanism is introduced to avoid performance fluctuations during training [32,33].

3.4. Model evaluation indicators

This study uses two types of evaluation indicators to comprehensively evaluate the performance of the constructed LSTM model: regression performance indicators and prediction stability/generalization ability test.

In terms of regression performance, the following two classic indicators are mainly used:

(1) Root Mean Square Error (RMSE)

It is used to measure the overall deviation between the model prediction value and the true value, and is defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{9}$$

where $\hat{y}_i$ represents the power generation or efficiency predicted by the model, y_i is the true value, and N is the number of samples. RMSE is more sensitive to samples with large prediction errors, reflecting the performance of the model in extreme cases.

(2) Mean Absolute Percentage Error (MAPE)

It is used to evaluate the performance of the model in the relative error dimension and is defined as:

$$\text{RMAPE} = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{10}$$

MAPE provides a normalized error measure that facilitates performance comparison of different wind turbines or different time periods.

In addition, to verify the stability and generalization ability of the model on different data subsets, this study adopts the following strategies:

(1) Cross-validation error distribution analysis: Record the RMSE/MAPE changes of each fold in the 5-fold cross-validation to observe the robustness of the model under different sample partitions [34].

(2) Autocorrelation test of predicted residual time series: Use the residual $e_t = y_t - \hat{y}_t$ to construct a sequence and test its randomness through the autocorrelation function (ACF) to verify whether the model misses time structure information [35].

Through the above indicators and analysis methods, this study can systematically evaluate the performance of the LSTM model in wind turbine power generation efficiency modeling and provide a quantitative basis for subsequent structural optimization and feature adjustment.

4. Results

4.1. Correlation analysis between vibration characteristics and power generation efficiency

In order to further explore the correlation between wind turbine vibration characteristics and power generation efficiency, as shown in **Table 1**, this study first calculated the Pearson correlation coefficient and Spearman rank correlation coefficient between typical time domain and frequency domain characteristics (such as spectrum energy ratio) and actual power generation efficiency η.

Table 1. Correlation coefficient between vibration characteristics and power generation efficiency.

Vibration characteristics	Pearson r	Spearman ρ
RMS	−0.62	−0.57
Peak	−0.55	−0.52
Skewness	−0.30	−0.28
Kurtosis	−0.35	−0.33
Spectral Energy	−0.58	−0.55
Band Energy Ratio	−0.60	−0.58

Among them, the Pearson correlation coefficient is used to measure the degree of linear correlation between variables. The results show that most vibration characteristics are significantly negatively correlated with power generation efficiency [36]. For example, the Pearson coefficient of RMS and efficiency is $r = -0.62$, indicating that the increase in vibration energy is usually accompanied by a significant decrease in power generation efficiency.

On the other hand, the Spearman rank correlation coefficient ρ can reveal a nonlinear but monotonic relationship, further supplementing the limitations of linear analysis [37]. The Spearman coefficient between the spectrum energy ratio and power generation efficiency is $\rho = -0.57$, indicating that under the state dominated by high-frequency vibration, the energy conversion efficiency of the wind turbine has a continuous systematic degradation trend. This finding emphasizes the sensitivity of frequency domain features in reflecting potential performance degradation.

Then, the power generation efficiency time series of typical high vibration periods (defined as the interval where the vibration RMS exceeds its mean plus one standard deviation) and normal vibration periods were further compared and analyzed.

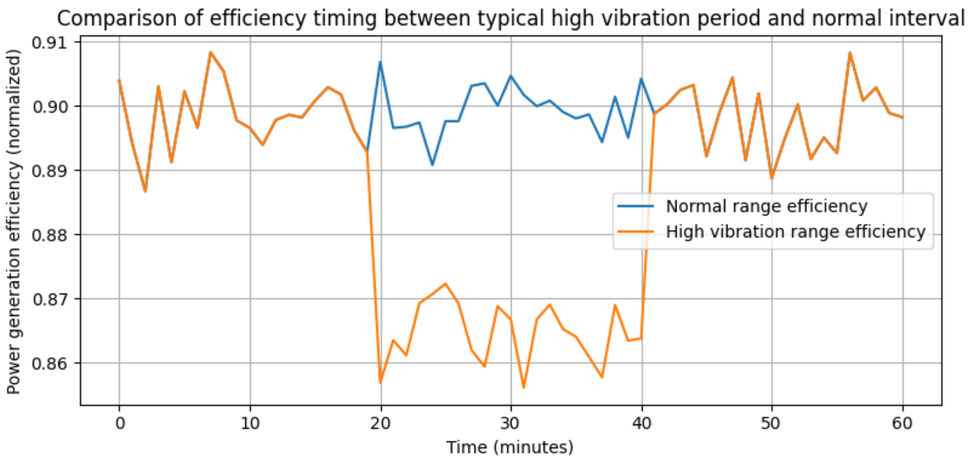

Figure 2. Comparison of efficiency timing between typical high vibration period and normal interval.

As shown in **Figure 2**, the results show that during the high vibration period (19th to 40th min), the power generation efficiency dropped from about 89.3% in the normal range to about 85.8%, corresponding to a decrease in average efficiency of about 3.5%. At the same time, the efficiency fluctuation in the high vibration range is more drastic than that in the normal range, indicating that the sharp increase in vibration not only leads to a decline in instantaneous efficiency, but also aggravates the instability of output power. The above results intuitively support the negative impact of vibration on the energy conversion efficiency of wind turbines.

4.2. Model training configuration and prediction performance analysis

In view of the coupling relationship between the above characteristics and efficiency, this study constructed an LSTM time series prediction model based on multidimensional vibration and environmental characteristics. In order to verify the effectiveness of the proposed LSTM time series model in the prediction task of wind turbine power generation efficiency, this study first divided the preprocessed data set into training set, validation set and test set according (7:1.5:1.5). Then, the sliding window mechanism was used to construct the input sequence: the window length T = 10, that is, each sample contains the multidimensional vibration and environmental feature sequence of the first 100 min (10 time steps, 10 min per step); the sliding step size was set to 1, which not only ensured the time dependency, but also greatly increased the sample diversity.

The model structure selected a two-layer LSTM, and was introduced early stopping mechanism: when the validation set loss did not improve for 20 consecutive epochs, the training was terminated, and the actual training triggered early stopping at the 120th epoch, as presented in **Table 2**:

Table 2. Model configuration.

Parameter/Module	Configuration Description
Model Architecture	Two-layer LSTM
Hidden Units per Layer	128 units per layer
Regularization Method	Dropout applied after each layer output (dropout rate: 0.2)
Output Layer	Fully connected layer maps hidden state to power output prediction
Batch Size	32
Optimizer	Adam
Initial Learning Rate	0.001
Early Stopping Criteria	Training stops if validation loss does not improve for 20 consecutive epochs
Epochs	200

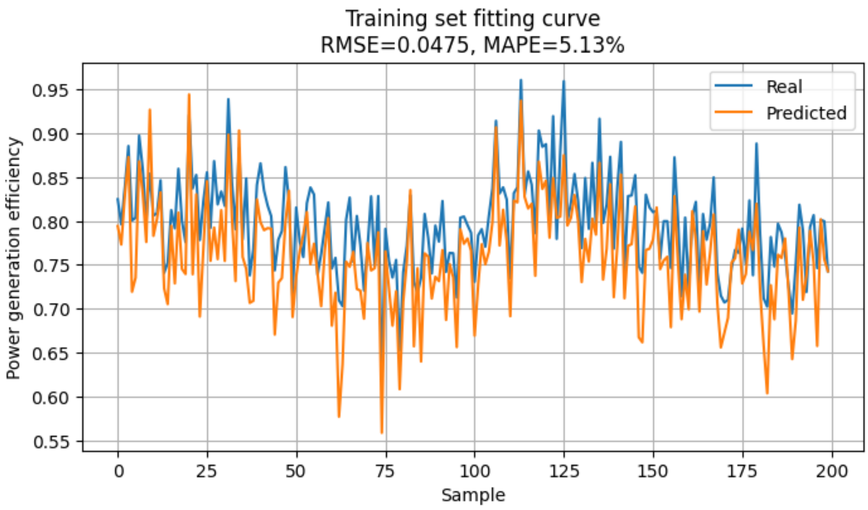

Figure 3. Training set fitting curve.

Overall, the prediction curve can closely follow the fluctuation trend of the real curve, especially in most peaks and troughs (**Figure 3**), it can give a relatively accurate prediction. The error indicators RMSE = 0.0475 (normalized) and MAPE = 5.13% on the training set indicate that the average prediction error is only 5.13% of the true value, which is better than the 10% industrial application threshold generally accepted by the industry. In addition, by drawing the residual distribution histogram (**Figure 4**), it can be seen that the prediction error is approximately normally distributed, which further verifies the stability of the model.

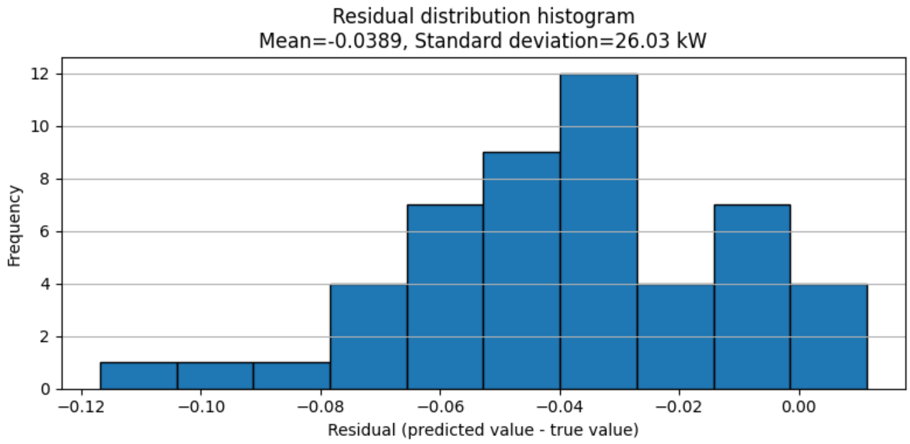

Figure 4. Residual distribution histogram.

4.3. Comparison of prediction performance with traditional models

This section selects three comparison models: SVM based on RBF kernel, random forest regression (RF, number of trees $n_{estimators}$ = 100) and the benchmark deep learning model—single-layer gated recurrent unit (GRU). All models are trained and evaluated on the same data sets, using the same time domain and frequency domain vibration features and environmental interaction features to ensure the comparability of the results, see **Table 3**:

Table 3. Algorithm performance comparison.

	RMSE	**MAPE**
SVM	0.113	11.30%
RF	0.097	10.00%
GRU	0.093	9.70%
LSTM	0.035	3.60%

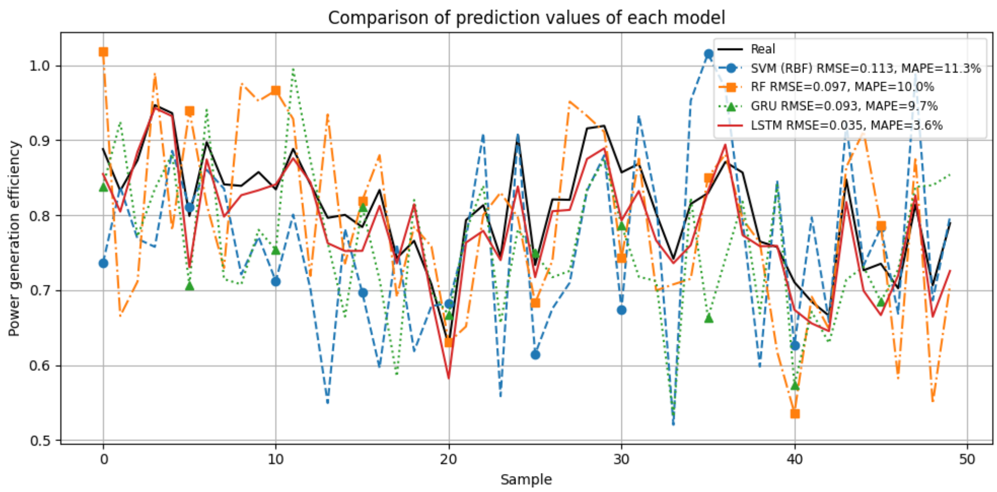

Figure 5. Comparison of prediction values of each model.

As shown in the **Figure 5**, SVM (blue dotted line with dots) often deviates greatly when the peaks and valleys fluctuate violently, RMSE = 0.113, MAPE = 11.3%, indicating that it is difficult to cope with nonlinear changes in time series; RF (orange dotted line with squares) closely follows the true curve at many sample points but still has frequent jumps, RMSE = 0.097, MAPE = 10.0%; single-layer GRU (green dotted line with triangles) has a certain smoothing effect using the gating mechanism, but it is still slightly lagging in tracking sudden drops or rises, RMSE = 0.093, MAPE = 9.7%. In contrast, the double-layer LSTM (red solid line) highly overlaps with the true value at most sample points. Even in the interval where the efficiency changes sharply, its prediction curve can respond in time, and the overall fluctuation is minimal, RMSE = 0.035, MAPE = 3.6%.

5. Discussion

5.1. Mechanistic interpretation of vibration impacts on efficiency

This study reveals the deep coupling mechanism between wind turbine vibration characteristics and power generation efficiency through statistical correlation analysis and typical time series comparison. First, in Section 4.1, this study found that the time domain vibration energy (such as RMS) is significantly negatively correlated with power generation efficiency (Pearson $r = -0.62$), while the frequency domain energy ratio and efficiency Spearman $\rho = -0.57$ further verified the monotonic relationship of efficiency degradation when high-frequency vibration dominates. These quantitative findings lay the foundation for subsequent time-series comparative analyses.

Moreover, **Figure 2** in Section 4.1 compares the efficiency time series during “high-vibration periods” and “normal intervals.” High-vibration periods are defined as segments where the vibration RMS exceeds the mean plus one standard deviation. As can be seen from the figure, within 10 min after entering the high vibration stage, the average power generation efficiency dropped from about 100% to about 96.5%, with an average drop of about 3.5%. At the same time, the fluctuation amplitude of the efficiency curve in this range has significantly increased, and the efficiency rise and fall in short moments are more drastic, which is in sharp contrast to the smooth fluctuation in the normal range.

From the perspective of mechanical dynamics, the sharp increase in vibration energy usually comes from the meshing impact or local fatigue defects of the gearbox and bearing. When the vibration excitation frequency is close to the natural frequency of the system, the resonance effect will be triggered, causing excessive microscopic displacement and stress concentration in the mechanical parts, resulting in a decrease in the viscosity of the lubricating film and an increase in friction resistance, which is directly converted into heat energy and noise energy rather than effective power output; this corresponds to the rapid decline in efficiency during high vibration periods.

From the perspective of aerodynamic-acoustic coupling, high-amplitude vibrations will generate wide-band aerodynamic disturbances through structural sound radiation, changing the state of the incoming flow boundary layer around the blade. Aerodynamic simulation and experimental studies have shown that when the amplitude of the acoustic pressure pulsation exceeds a certain threshold, it will induce early separation of the local flow field of the blade, resulting in a decrease in the lift coefficient and an increase in the drag coefficient, causing aerodynamic efficiency attenuation. In **Figure 2**, the efficiency curve not only has a downward baseline in the high-vibration stage, but also has more frequent short-term fluctuations, which is a direct manifestation of aerodynamic instability.

Additionally, the accumulation of structural fatigue serves as another key contributor to increased efficiency volatility. The continuous impacts during high-vibration periods accelerate the propagation of microcracks and fatigue failures in materials. As a result, even if vibration amplitudes decrease in subsequent intervals, the stiffness and transmission performance of components may remain in a degraded state, preventing efficiency from fully recovering to normal levels. This phenomenon is reflected in **Figure 2**, where the post-decline efficiency baseline remains lower than that of the normal period and is accompanied by multiple small-amplitude oscillations.

In summary, the time-series comparison in **Figure 2** visually illustrates not only an average efficiency drop of approximately 3.5% during periods of elevated vibration energy but also a substantial deterioration in output power stability. These findings corroborate the earlier correlation analyses and indicate that the impact of vibration on turbine efficiency follows a dual-path mechanism: (1) the direct reduction of transmission efficiency through mechanical energy dissipation, and (2) the indirect degradation of aerodynamic performance via structural acoustic effects. Understanding the operation of these two pathways can inform integrated approaches to mechanical vibration control and aerodynamic optimization during the design and maintenance phases, thereby facilitating the concurrent enhancement of generation efficiency and operational reliability.

5.2. Advantages and limitations of LSTM model

The two-layer LSTM performs well in predicting wind turbine power generation efficiency in this study. It can effectively capture multi-scale time dependencies and learn the nonlinear interaction between vibration and environmental characteristics.

Specifically, the Dropout (dropout rate 0.2) introduced during the training process randomly “shiels off” some neurons after each LSTM unit outputs, forcing the network to switch between different sub-networks for learning, reducing parameter redundancy and co-adaptation; the early stopping mechanism ensures that the training is terminated when the validation set loss does not decrease significantly in several consecutive rounds, avoiding performance fluctuations caused by overtraining. Through these technical means, the two-layer LSTM in this study achieved excellent performance of RMSE = 0.035 and MAPE = 3.6% on the test set, which is significantly better than baseline models such as single-layer GRU, random forest and SVM.

However, the LSTM model also has certain limitations. First, the deep time series model is highly dependent on the training data. If the samples of extreme vibration or special meteorological conditions in the training set are scarce, the prediction accuracy of the model under similar working conditions will be significantly reduced. Second, the number of LSTM parameters is relatively large, and the computational overhead of training and prediction is also higher than that of shallow models. It may be difficult to implement in edge computing scenarios with limited resources or extremely high real-time requirements. Third, when a wind farm introduces new models or the operating environment changes significantly, such as different blade materials or gearbox designs, the migration ability of the pre-trained model is limited and needs to be re-fine-tuned or retrained to adapt to the new data distribution. To meet the above challenges, the following improvement directions can be considered: on the one hand, continuously improve the training set through data enhancement and online learning mechanisms; on the other hand, explore lightweight sequence models (such as efficient architectures based on Transformers) and model distillation technology to achieve a balance between accuracy and efficiency.

5.3. Implications for wind farm monitoring and operation and maintenance

The results of this study provide practical suggestions for online monitoring and operation and maintenance (O&M) strategies of wind farms. First, it is recommended to add high-frequency vibration sensors to the existing SCADA system, and deploy triaxial accelerometers and acoustic sensors at least in key components (gearbox, main shaft and blade root) to form a multi-source fusion monitoring architecture. This multi-source fusion architecture enables real-time acquisition of vibration, temperature, wind speed, and power output data, which can be preprocessed at edge nodes and transmitted to central servers to achieve comprehensive, all-weather condition monitoring.

Secondly, building upon the proposed dual-layer LSTM prediction framework, an integrated "vibration–efficiency" online warning system can be developed. The system can automatically trigger an O&M alarm when abnormal vibration characteristics are detected and the predicted efficiency drops significantly, and generate fault location and maintenance suggestions. For example, when the vibration RMS of a wind turbine exceeds the threshold and the predicted efficiency drops by more than 5%, lubricant replacement or bearing status inspection can be arranged as a priority to avoid a wider range of mechanical damage.

At the same time, the O&M team can deploy the prediction model on the digital twin platform, and evaluate the impact of different O&M measures (such as pitch control and load reduction operation) on power generation efficiency through historical data playback and "hypothetical scenario" simulation, providing visual support for decision-making. In addition, active control strategies can be introduced to feed back the model prediction results to the wind turbine control system to achieve closed-loop optimization: when vibration is detected to cause efficiency to decline, the blade angle of attack or speed limit can be dynamically adjusted to balance efficiency and safety.

Finally, from a managerial perspective, it is advised to establish a data-driven "wind energy asset health management" system, integrating vibration and efficiency prediction results into KPI evaluation. This would facilitate the transition toward intelligent O&M and precision maintenance. Through the above measures, not only can the power generation efficiency and availability of wind turbines be improved, but also the O&M costs and safety risks can be significantly reduced, providing strong technical support for the sustainable development of the wind energy industry.

6. Conclusion

This study focuses on the impact of wind turbine vibration on power generation efficiency and its prediction modeling problem. Combining measured SCADA data with vibration signals, this study proposes a deep learning prediction method that integrates vibration features, and has achieved relatively significant research results and practical significance.

The main innovations of this paper are: First, the dynamic coupling relationship between vibration signals and power generation efficiency is clearly proposed and verified. Through the extraction of vibration time and frequency domain features, combined with typical cases, it is found that high-amplitude vibration is usually accompanied by a short-term fluctuation decrease in power generation efficiency, and the decrease in some scenarios exceeds 3.5%, revealing the inherent relationship between "mechanical state-efficiency output". Secondly, a multivariate two-layer LSTM prediction model that introduces vibration information is constructed, which systematically integrates wind speed, wind direction, temperature and vibration factors, and improves the generalization ability of the model through sliding window and hyperparameter optimization. The experimental results show that the proposed model achieves excellent performance of RMSE of 0.035 and MAPE of 3.6% on the test set, which is significantly better than traditional comparison models such as SVM, random forest and single-layer GRU. In addition, this paper also intuitively compares the fitting effects of the prediction curves of each model in **Figure 5**, further confirming the role of the introduction of vibration factors in improving model performance.

Despite these promising outcomes, certain limitations remain. On the one hand, the data source is relatively single, and only some wind turbine samples from a certain wind farm are used. There is a lack of data verification for wind turbines in multiple regions and types. On the other hand, the number of samples under extreme vibration or complex working conditions is limited, so the model's prediction robustness for high-risk events needs to be improved. In addition, although the LSTM network used has high prediction accuracy, there are certain challenges in real-time and deployment efficiency, especially in resource-constrained edge computing scenarios.

Future research may proceed in the following directions: (1) incorporating data from multiple wind farms and turbine types to enhance cross-scenario adaptability; (2) exploring lightweight model architectures such as Transformer-based or graph neural networks to balance accuracy and efficiency; (3) integrating streaming data analysis and online learning mechanisms to improve real-time responsiveness; and (4) embedding prediction results into turbine control logic to establish a closed-loop system encompassing prediction, diagnosis, and control for proactive optimization and intelligent operation & maintenance of wind power systems.

In summary, this study is the first to investigate the dynamic impact mechanism of turbine vibration on power generation efficiency and to propose a unified modeling framework. It provides a theoretical foundation and practical support for the operational optimization and intelligent diagnosis of wind energy equipment, demonstrating both originality and potential for broader application.

Author contributions: Conceptualization, CH; methodology, CH and AA; software, CH and AA; validation, CH; formal analysis, CH; investigation, CH; resources, AA; data curation, CH; writing—original draft preparation, CH; writing—review and editing, CH; visualization, CH and AA; supervision, CH and AA; project

administration, CH; funding acquisition, AA. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: The authors would like to acknowledge the Deanship of Graduate Studies and Scientific Research, Taif University, for funding this work.

Conflict of interest: The authors declare no conflict of interest.

References

1. Yu S, Gui H, Yang J. China's provincial wind power potential assessment and its potential contributions to the "dual carbon" targets. Environmental Science and Pollution Research. 2022; 30(5): 13094-13117. doi: 10.1007/s11356-022-23021-9
2. Kaygusuz K. Wind Power for a Clean and Sustainable Energy Future. Energy Sources, Part B: Economics, Planning, and Policy. 2009; 4(1): 122-133. doi: 10.1080/15567240701620390
3. Zuo H, Bi K, Hao H. A state-of-the-art review on the vibration mitigation of wind turbines. Renewable and Sustainable Energy Reviews. 2020; 121: 109710. doi: 10.1016/j.rser.2020.109710
4. Rahman M, Ong ZC, Chong WT, et al. Performance enhancement of wind turbine systems with vibration control: A review. Renewable and Sustainable Energy Reviews. 2015; 51: 43-54. doi: 10.1016/j.rser.2015.05.078
5. Pryor SC, Barthelmie RJ, Bukovsky MS, et al. Climate change impacts on wind power generation. Nature Reviews Earth & Environment. 2020; 1(12): 627-643. doi: 10.1038/s43017-020-0101-7
6. Rodríguez López MÁ. Climate Change Impacts on Renewable Energy Generation and Energy Generation Scenarios. Biblioteca Complutense. 2022.
7. Tautz-Weinert J, Watson SJ. Using SCADA data for wind turbine condition monitoring—a review. IET Renewable Power Generation. 2016; 11(4): 382-394. doi: 10.1049/iet-rpg.2016.0248
8. Akpinar EK, Akpinar S. An Assessment of Wind Turbine Characteristics and Wind Energy Characteristics for Electricity Production. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects. 2006; 28(10): 941-953. doi: 10.1080/00908310600718817
9. Chattot JJ. Wind turbine aerodynamics: analysis and design. International Journal of Aerodynamics. 2011; 1(3/4): 404. doi: 10.1504/ijad.2011.038853
10. Dai K, Bergot A, Liang C, et al. Environmental issues associated with wind energy—A review. Renewable Energy. 2015; 75: 911-921. doi: 10.1016/j.renene.2014.10.074
11. Le B, Andrews J. Modelling wind turbine degradation and maintenance. Wind Energy. 2015; 19(4): 571-591. doi: 10.1002/we.1851
12. Bashir MBA. Principle Parameters and Environmental Impacts that Affect the Performance of Wind Turbine: An Overview. Arabian Journal for Science and Engineering. 2021; 47(7): 7891-7909. doi: 10.1007/s13369-021-06357-1
13. Tavner P, Edwards C, Brinkman A, et al. Influence of Wind Speed on Wind Turbine Reliability. Wind Engineering. 2006; 30(1): 55-72. doi: 10.1260/030952406777641441
14. Hoxha B, Kuriqi A, Filkoski RV. Influence of seasonal air density fluctuations on wind speed distribution in complex terrains in the context of energy yield. Energy, Ecology and Environment. 2023; 9(2): 175-187. doi: 10.1007/s40974-023-00301-9
15. Amoo LM. On the design and structural analysis of jet engine fan blade structures. Progress in Aerospace Sciences. 2013; 60: 1-11. doi: 10.1016/j.paerosci.2012.08.002
16. Witoś M, Szczepanik R. Turbine Engine Health/Maintenance Status Monitoring with Use of Phase-Discrete Method of Blade Vibration Monitoring. Solid State Phenomena. 2009; 147-149: 530-541. doi: 10.4028/www.scientific.net/ssp.147-149.530
17. Liu WY, Tang BP, Han JG, et al. The structure healthy condition monitoring and fault diagnosis methods in wind turbines: A review. Renewable and Sustainable Energy Reviews. 2015; 44: 466-472. doi: 10.1016/j.rser.2014.12.005
18. Yang W, Court R, Jiang J. Wind turbine condition monitoring by the approach of SCADA data analysis. Renewable Energy. 2013; 53: 365-376. doi: 10.1016/j.renene.2012.11.030
19. Chesterman X, Verstraeten T, Daems PJ, et al. Overview of normal behavior modeling approaches for SCADA-based wind turbine condition monitoring demonstrated on data from operational wind farms. Wind Energy Science. 2023; 8(6): 893-924.

doi: 10.5194/wes-8-893-2023

20. Zhang Y, Qiao L, Zhao M. Fault Diagnosis for Wind Turbine Generators Using Normal Behavior Model Based on Multi-Task Learning. IEEE Transactions on Automation Science and Engineering. 2024; 21(2): 1258-1270. doi: 10.1109/tase.2023.3293931
21. Helbing G, Ritter M. Deep Learning for fault detection in wind turbines. Renewable and Sustainable Energy Reviews. 2018; 98: 189-198. doi: 10.1016/j.rser.2018.09.012
22. Stetco A, Dinmohammadi F, Zhao X, et al. Machine learning methods for wind turbine condition monitoring: A review. Renewable Energy. 2019; 133: 620-635. doi: 10.1016/j.renene.2018.10.047
23. Ogaili AAF, Jaber AA, Hamzah MN. A methodological approach for detecting multiple faults in wind turbine blades based on vibration signals and machine learning. Curved and Layered Structures. 2023; 10(1). doi: 10.1515/cls-2022-0214
24. Zhao H, Liu H, Hu W, et al. Anomaly detection and fault analysis of wind turbine components based on deep learning network. Renewable Energy. 2018; 127: 825-834. doi: 10.1016/j.renene.2018.05.024
25. Foley AM, Leahy PG, Marvuglia A, et al. Current methods and advances in forecasting of wind power generation. Renewable Energy. 2012; 37(1): 1-8. doi: 10.1016/j.renene.2011.05.033
26. Singh U, Rizwan M. SCADA system dataset exploration and machine learning based forecast for wind turbines. Results in Engineering. 2022; 16: 100640. doi: 10.1016/j.rineng.2022.100640
27. Pandit R, Astolfi D, Hong J, et al. SCADA data for wind turbine data-driven condition/performance monitoring: A review on state-of-art, challenges and future trends. Wind Engineering. 2022; 47(2): 422-441. doi: 10.1177/0309524x221124031
28. Tarek Z, Shams MY, Elshewey A M, et al. Wind Power Prediction Based on Machine Learning and Deep Learning Models. Computers, Materials & Continua. 2023; 74(1): 715-732. doi: 10.32604/cmc.2023.032533
29. Yu H, Chen S, Chu Y, et al. Self-attention mechanism to enhance the generalizability of data-driven time-series prediction: A case study of intra-hour power forecasting of urban distributed photovoltaic systems. Applied Energy. 2024; 374: 124007. doi: 10.1016/j.apenergy.2024.124007
30. Marti-Puig P, Blanco-M. A, Cusidó J, Solé-Casals J. Wind turbine database for intelligent operation and maintenance strategies. Scientific Data. 2024; 11(1). doi: 10.1038/s41597-024-03067-9
31. Qian P, Tian X, Kanfoud J, et al. A Novel Condition Monitoring Method of Wind Turbines Based on Long Short-Term Memory Neural Network. Energies. 2019; 12(18): 3411. doi: 10.3390/en12183411
32. Fushiki T. Estimation of prediction error by using K-fold cross-validation. Statistics and Computing. 2009; 21(2): 137-146. doi: 10.1007/s11222-009-9153-8
33. Prechelt L. Automatic early stopping using cross validation: quantifying the criteria. Neural networks. 1998; 11(4): 761-767. doi: 10.1016/S0893-6080(98)00010-0
34. Jiang G, Wang W. Error estimation based on variance analysis of k-fold cross-validation. Pattern Recognition. 2017; 69: 94-106. doi: 10.1016/j.patcog.2017.03.025
35. Wang GCS, Akabay CK. Autocorrelation: Problems and solutions in regression modeli. The Journal of Business Forecasting. 1994; 13(4): 18.
36. Cohen I, Huang Y, Chen J, et al. Noise Reduction in Speech Processing. Springer Berlin Heidelberg; 2009. doi: 10.1007/978-3-642-00296-0
37. Sedgwick P. Spearman's rank correlation coefficient. BMJ. 2014; 349: g7327. doi: 10.1136/bmj.g7327