Mô hình LSTM dựa trên tối ưu hóa Bayes dùng cho dự báo ngắn hạn công suất điện gió

TÓM TẮT

Phần lớn các phương pháp dự báo công suất điện gió hiện nay chủ yếu dựa trên dữ liệu chuỗi thời gian lịch sử của công suất phát tua-bin gió. Tuy nhiên, quá trình phát điện gió chịu ảnh hưởng đáng kể bởi các thông số vận hành tua-bin, những yếu tố thường bị bỏ qua và có thể làm giảm độ chính xác dự báo. Trong nghiên cứu này, dữ liệu vận hành thu thập từ hệ thống SCADA của một tua-bin gió thực tế được sử dụng nhằm nâng cao hiệu quả dự báo công suất điện gió. Phương pháp hồi quy từng bước được áp dụng để xác định các thông số vận hành có ảnh hưởng quan trọng đến công suất phát, qua đó giảm số chiều dữ liệu đầu vào và nâng cao khả năng diễn giải của mô hình. Mô hình thống kê truyền thống ARIMAX được xây dựng làm mô hình tham chiếu cho bài toán dự báo. Đồng thời, mô hình học sâu LSTM được triển khai nhằm mô tả đặc tính phi tuyến và sự biến thiên theo thời gian của dữ liệu công suất gió, trong đó các siêu tham số được tối ưu bằng phương pháp tối ưu hóa Bayes. Hiệu quả dự báo của mô hình LSTM sau tối ưu hóa được so sánh với mô hình ARIMAX. Kết quả thực nghiệm cho thấy mô hình LSTM dựa trên tối ưu hóa Bayes có kết quả dự báo tốt hơn so với mô hình ARIMAX theo các tiêu chí MAE và WAPE. Kết quả này khẳng định hiệu quả của việc kết hợp các thông số vận hành quan trọng và tối ưu hóa Bayes trong việc nâng cao độ chính xác dự báo công suất điện gió.

Từ khóa: *Dự báo điện gió, ARIMAX, Tối ưu hóa Bayes, LSTM.*

Bayesian optimization-based LSTM model for short-term wind power forecasting

ABSTRACT

Most existing wind power forecasting methods rely primarily on historical time-series data of wind turbine output. However, wind power generation is strongly influenced by turbine operating parameters, which are often neglected and may degrade forecasting accuracy. In this study, operational data collected from a practical wind turbine SCADA system are utilized to enhance wind power forecasting performance. First, stepwise regression is employed to identify the most significant operating parameters affecting wind turbine power output, thereby reducing input dimensionality and improving model interpretability. Next, a traditional statistical autoregressive integrated moving average with exogenous variables (ARIMAX) model is developed as a benchmark forecasting approach. In addition, a deep learning model, namely a long short-term memory (LSTM) network, is implemented to capture the nonlinear and temporal characteristics of wind power data. To further improve forecasting accuracy, the hyperparameters of the LSTM model are optimized using Bayesian optimization. The Bayesian optimization–based LSTM model is then proposed as an alternative approach for wind power forecasting. The forecasting performance of the optimized LSTM model is systematically compared with that of the ARIMAX model. Experimental results demonstrate that the optimized LSTM model significantly outperforms the traditional ARIMAX approach in terms of mean absolute error (MAE) and weighted absolute percentage error (WAPE). These findings confirm the effectiveness of incorporating significant operating parameters and Bayesian hyperparameter optimization in improving wind power forecasting accuracy.

Keywords: *Wind power forecasting, ARIMAX, Bayesian optimization, LSTM.*

1. INTRODUCTION

In recent decades, there has been a rapid expansion of renewable energy sources, among which wind power has emerged as one of the most prominent contributors to the global energy transition. Compared with solar power, wind energy offers several distinct advantages within modern power systems. Wind power generally exhibits a higher capacity factor, enabling more efficient utilization of installed generation capacity over long operating periods. Furthermore, wind turbines are capable of producing electricity continuously throughout both daytime and nighttime, whereas solar power generation is strictly constrained by solar irradiance availability.

In addition, the variability of wind power output is often more gradual and less sensitive to short-term atmospheric disturbances, such as passing cloud cover, than solar photovoltaic generation. This characteristic contributes to smoother power profiles and enhances short-term operational stability at the system level. Wind power also allows for flexible land use, particularly in onshore installations where agricultural or industrial activities can coexist with wind farms. Moreover, wind energy demonstrates strong potential for large-scale deployment, especially in offshore environments, where higher and more consistent wind speeds enable substantial generation capacity. As a result, wind power plays a crucial role in reducing greenhouse gas emissions, mitigating climate change, and promoting sustainable development. Owing to these advantages, wind power has been widely integrated into power systems worldwide and continues to experience rapid growth.

Despite its considerable benefits, wind power also presents significant challenges due to its inherent dependence on meteorological conditions. The electrical output of wind turbines varies nonlinearly with wind speed and is further influenced by wind direction, air density, and turbine operating states, leading to pronounced intermittency and uncertainty. These fluctuations introduce substantial difficulties for power system operators, including increased requirements for spinning reserves, frequency and voltage regulation, and real-time balancing actions. As the penetration level of wind power continues to rise, forecasting inaccuracies may exacerbate supply–demand imbalances, trigger local network congestion, compromise system reliability, and negatively affect electricity market operations.

Consequently, the development of accurate and reliable wind power forecasting techniques is of critical importance for generation planning, economic dispatch, reserve allocation, and the secure and efficient operation of modern power systems.

Fundamentally, wind power forecasting can be classified into several categories, including very short-term, short-term, medium-term, and long-term forecasting, corresponding to different objectives and requirements in power system operation and planning. Very short-term forecasting, typically ranging from several minutes to a few hours, mainly supports real-time control, frequency and voltage stability, and the integration of renewable energy sources into the grid. Short-term forecasting, covering time horizons from several hours to several days, plays an important role in unit commitment, generation dispatch, reserve management, and participation in electricity markets. Medium-term forecasting, usually extending from several days to several weeks, is used for maintenance planning, fuel management, and reliability assessment of power system operation. Meanwhile, long-term forecasting, with horizons ranging from several months to several years, primarily serves generation and transmission planning, investment evaluation, and the formulation of long-term energy policies [1].

To address the wind power forecasting problem, numerous forecasting methods have been proposed in the literature. These models can be broadly categorized into statistical models, artificial intelligence (AI)-based models, and hybrid models. Statistical models applied to wind power forecasting include exponential smoothing approaches [2,3], autoregressive (AR) models [4], autoregressive moving average (ARMA) models [5,6], and autoregressive integrated moving average (ARIMA) models [7,8]. Statistical models are among the earliest and most widely used approaches for wind power forecasting due to their simple structure, clear interpretability, and low computational cost. Traditional time-series models such as AR, MA, and ARIMA exploit linear relationships between wind power output and past values of the data series, often yielding satisfactory results for relatively stable time series with smooth variations [9]. However, the performance of statistical models strongly depends on assumptions of linearity and stationarity, as well as appropriate selection of input variables. Consequently, when applied to wind power systems with high variability and pronounced nonlinear characteristics, statistical models are typically used as benchmark references for comparison with more advanced machine learning and deep learning approaches [10].

In the field of wind power forecasting, numerous AI–based models have been proposed to capture nonlinear relationships and complex dynamics in time-series data. Feedforward neural networks, such as the multilayer perceptron [11], commonly trained using the back-propagation neural network (BP NN) algorithm [12], were among the earliest models applied and demonstrated improved performance over traditional statistical methods. To better exploit temporal dependencies, recurrent neural networks (RNNs) were developed, with variants such as the Elman neural network and layered RNNs used to retain past state information of time-series data [13]. In recent years, deep recurrent learning models such as long short-term memory (LSTM) [14] and gated recurrent unit networks [15] have become dominant approaches in wind power forecasting due to their ability to effectively capture long-term dependencies and mitigate the vanishing gradient problem. Variants such as bidirectional LSTM (BiLSTM) further exploit information in both temporal directions to improve forecasting accuracy [16]. In addition, the echo state network, a type of recurrent network with randomly generated reservoirs, has also been investigated as a computationally efficient solution for wind power time-series forecasting [17]. Beyond neural networks, machine learning methods such as support vector machines (SVM) [18] and gradient boosting regression trees [19] have been widely applied to wind power forecasting owing to their robustness in handling nonlinear and noisy data. Finally, ensemble models that combine multiple forecasting methods have been proposed to enhance both accuracy and robustness of forecasting results [20].

In recent years, combined models have been extensively studied in wind power forecasting to leverage the strengths of different methods while mitigating the limitations of individual models. These hybrid approaches often integrate signal processing techniques—such as empirical mode decomposition (EMD), variational mode decomposition (VMD), or wavelet transform—with forecasting models including neural networks, deep learning models, or machine learning algorithms, in order to decompose wind power series into different oscillatory components and improve the learning capability of forecasting models [21,22]. Other hybrid approaches, including autoregressive fractionally integrated moving average combined with least squares support vector machines [23], boosting algorithms combined with ARMA models [24], hybrid CEEMDAN–EWT deep learning methods [25], as well as neuro-wavelet and LSTM models, have demonstrated improved forecasting accuracy and robustness under highly variable wind conditions. Empirical studies indicate that hybrid models generally outperform single models, particularly in short-term and very short-term wind power forecasting [26].

Nevertheless, most existing models still primarily rely on wind power time-series data, while the in-depth exploitation of detailed wind turbine operational parameters obtained from SCADA systems remains limited and has not been comprehensively investigated. Recent studies have begun to incorporate wind turbine operational parameters, for example by applying response surface methodology, in which variables such as wind speed, nacelle position, pitch angle, and ambient temperature are used to forecast turbine power output [27]. Recurrent neural network–based models, including nonlinear autoregressive neural networks with external inputs, layer recurrent neural network models, distributed delay neural network models, and time delay neural network models, have also utilized operational parameters such as wind speed, pitch angle, ambient temperature, nacelle position, and wind direction, in addition to turbine power output, to improve forecasting performance [28].

Despite the extensive body of research on wind power forecasting, several fundamental limitations remain insufficiently addressed. Most existing studies primarily exploit historical wind power time-series data, often supplemented by a limited number of meteorological variables, while the rich and high-resolution operational information available from wind turbine SCADA systems is largely underutilized. Moreover, input variable selection in many previous works is commonly based on empirical assumptions or prior experience, lacking a systematic and statistically grounded procedure to identify the most influential operational parameters, which may introduce input redundancy, obscure physical interpretability, and degrade model generalization. Furthermore, although deep learning models—particularly LSTM networks—have demonstrated strong potential in capturing nonlinear temporal dependencies, their forecasting performance is highly sensitive to hyperparameter selection. However, hyperparameter tuning is often conducted using trial-and-error or grid search strategies, which are computationally inefficient and prone to suboptimal solutions.

To overcome these limitations, this study leverages high-resolution SCADA data from a real operating wind turbine, incorporating a comprehensive set of operational parameters. A stepwise regression method is first employed to rigorously identify the statistically significant variables that most strongly influence wind turbine power output, thereby enhancing model interpretability and reducing input dimensionality. An LSTM-based forecasting model is then developed to effectively capture the nonlinear and dynamic characteristics of wind power generation. A key contribution of this work is the integration of Bayesian optimization (BO) for the systematic and efficient tuning of LSTM hyperparameters, enabling the model to achieve improved predictive performance while avoiding excessive computational cost. The proposed approach is benchmarked against a conventional ARIMAX model, and comparative results demonstrate that the proposed method consistently delivers superior forecasting accuracy, as evidenced by lower mean absolute error (MAE) and weighted absolute percentage error (WAPE). An overview of the proposed wind power forecasting approach is demonstrated in Figure 1.

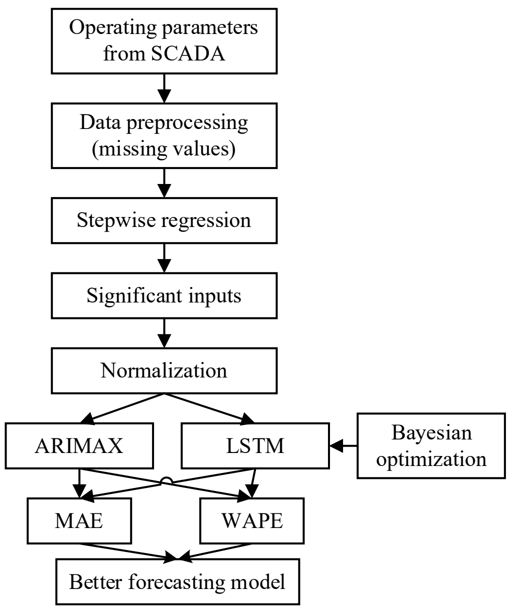

Figure 1. Overview of the proposed wind power forecasting approach.

2. METHODOLOGY

2.1. Stepwise regression

Stepwise regression is a statistical feature selection approach built upon the multiple linear regression framework, aiming to identify a parsimonious subset of explanatory variables that significantly affect the response variable. The relationship between the dependent variable and candidate predictors is expressed as

$$y = \beta_0 + \sum_{i=1}^{m} \beta_i x_i + \varepsilon \tag{1}$$

where y denotes the response variable (e.g., wind turbine power output), x_i (i = 1, …, m) represent candidate explanatory variables (e.g., operating parameters obtained from the SCADA system), β_0 is the intercept, β_i are regression coefficients, and ε is a random error term.

The stepwise procedure iteratively updates the regression model by adding or removing variables based on their statistical significance. A commonly used criterion is the p-value, which represents the probability of observing a regression coefficient at least as extreme as the estimated one under the null hypothesis that the coefficient equals zero. Variables with p-values below a predefined significance level (e.g., 0.05) are retained, while statistically insignificant variables are removed. In addition, model selection can be guided by information-theoretic criteria such as the Akaike information criterion (AIC), defined as

$$AIC = n \ln\left(\frac{RSS}{n}\right) + 2m \tag{2}$$

where RSS is the residual sum of squares of the regression model, n is the number of observations, and m denotes the number of explanatory variables included in the model. Lower values of AIC indicate a better trade-off between model goodness-of-fit and complexity.

This iterative process continues until no further improvement can be achieved according to the selected criteria. In wind power forecasting applications, stepwise regression provides a statistically grounded and interpretable method for identifying influential operating parameters from high-dimensional SCADA data, thereby reducing input redundancy and enhancing the robustness and generalization capability of subsequent forecasting models.

2.2. ARIMAX model

The autoregressive integrated moving average with exogenous variables (ARIMAX) model is a classical statistical method for time-series forecasting that extends the ARIMA framework by incorporating external explanatory variables. In wind power forecasting, ARIMAX enables the inclusion of selected operating or environmental parameters while modeling the linear temporal dependence inherent in wind power time-series data.

Let y_t denote the wind power output at time t, and let $x_{k,t}$ (k = 1, …, K) represent the exogenous variables. The ARIMAX (p, d, q) model in explicit time-domain form is expressed as:

$$\begin{aligned} \Delta^d y_t = c + \sum_{i=1}^{p} \phi_i \Delta^d y_{t-i} \\ + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{K} \gamma_k x_{k,t} + \varepsilon_t \end{aligned} \tag{3}$$

where Δ^d is the differencing operator of order d, c is a constant term, ϕ_i and θ_j are the autoregressive and moving average coefficients, respectively, γ_k denotes the regression coefficient associated with the k^{th} exogenous variable, and ε_t represents a zero-mean white noise process.

The selection of the ARIMAX model order (p, d, q) is commonly performed using the Bayesian information criterion (BIC), which provides a trade-off between model accuracy and complexity. The BIC is defined as

$$BIC = n \ln(\sigma^2) + r \ln(n) \tag{4}$$

where n is the number of observations, σ^2 is the estimated variance of the residuals, and r denotes the total number of estimated parameters in the model. A lower BIC value indicates a more parsimonious and statistically preferable model. Owing to its interpretability and well-established theoretical foundation, the ARIMAX model is frequently employed as a benchmark for evaluating advanced machine learning and deep learning–based wind power forecasting methods.

2.3. LSTM network

LSTM networks are a special class of RNNs designed to effectively model long-term temporal dependencies in sequential data. Unlike conventional RNNs, which often suffer from vanishing or exploding gradient problems, LSTM introduces a memory cell and gating mechanisms that regulate information flow over time. Owing to these characteristics, LSTM has been widely applied in wind power forecasting to capture nonlinear and time-dependent patterns in wind power time-series data.

An LSTM unit consists of a cell state and three main gates: the forget gate, input gate, and output gate. Given an input vector $\mathbf{x}_t$ at time step *t*, the hidden state $\mathbf{h}_{t-1}$, and the cell state $\mathbf{c}_{t-1}$, the LSTM operations are defined as follows:

Forget gate

$$\mathbf{f}_t = \sigma\left(\mathbf{W}_f\mathbf{x}_t + \mathbf{U}_f\mathbf{h}_{t-1} + \mathbf{b}_f\right) \tag{5}$$

Input gate

$$\mathbf{i}_t = \sigma\left(\mathbf{W}_i\mathbf{x}_t + \mathbf{U}_i\mathbf{h}_{t-1} + \mathbf{b}_i\right) \tag{6}$$

Candidate cell state

$$\tilde{\mathbf{c}}_t = \tanh\left(\mathbf{W}_c\mathbf{x}_t + \mathbf{U}_c\mathbf{h}_{t-1} + \mathbf{b}_c\right) \tag{7}$$

Cell state update

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t \tag{8}$$

Output gate

$$\mathbf{o}_t = \sigma\left(\mathbf{W}_o\mathbf{x}_t + \mathbf{U}_o\mathbf{h}_{t-1} + \mathbf{b}_o\right) \tag{9}$$

Hidden state

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t) \tag{10}$$

where σ denotes the sigmoid activation function, $\tanh(\cdot)$ is the hyperbolic tangent function, $\odot$ represents element-wise multiplication, $\mathbf{W}$ and $\mathbf{U}$ are weight matrices, and $\mathbf{b}$ are bias vectors.

In wind power forecasting applications, the LSTM network learns a nonlinear mapping between historical wind power observations (and optionally additional input variables) and future power output. By maintaining a memory of relevant past information, LSTM is capable of modeling complex temporal dependencies and sudden power fluctuations caused by changing wind conditions.

2.4. BO for LSTM Hyperparameter Tuning

The forecasting performance of LSTM networks is highly sensitive to hyperparameter selection. Conventional tuning strategies, such as manual adjustment or grid search, are computationally inefficient and often fail to identify globally optimal configurations, especially when model training is costly. To overcome these limitations, BO is adopted to systematically determine the optimal hyperparameters for the proposed LSTM-based wind power forecasting model.

2.4.1. Mathematic formulas

Let y_t and y_t denote the actual and predicted wind power output at time step *t*, espectively. The forecasting performance is evaluated using the mean absolute error (MAE) and weighted absolute percentage error (WAPE), defined as:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|y_t - y_t\right| \tag{11}$$

$$WAPE = \frac{\sum_{t=1}^{N}\left|y_t - y_t\right|}{\sum_{t=1}^{N}\left|y_t\right|} \times 100\% \tag{12}$$

where N represents the total number of validation samples.

It should be noted that the purpose of this study is comparative model evaluation under a unified supervised learning structure. The ARIMAX model is employed as the conventional statistical baseline, and no separate naïve short-horizon benchmark is considered.

To jointly account for absolute and relative forecasting errors, a composite objective function is constructed as

$$J = \alpha MAE + (1-\alpha)WAPE \tag{13}$$

where $\alpha \in [0,1]$ is a weighting coefficient.

The hyperparameter optimization problem is formulated as

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}\in\Omega} f(\boldsymbol{\theta}) \tag{14}$$

where $f(\boldsymbol{\theta})$ corresponds to MAE or WAPE, and Ω denotes the feasible hyperparameter search space.

2.4.2. Hyperparameter search space

The vector of LSTM hyperparameters to be optimized is defined as

$$\boldsymbol{\theta} = [N_h, \eta, N_e, d, L] \tag{15}$$

where N_h is the number of hidden units in the LSTM layer, η is the learning rate, N_e is the number of training epochs, d is the dropout rate applied to mitigate overfitting, L is the look-back window length representing the number of historical time steps used as input.

These hyperparameters are selected due to their significant influence on the model's learning capacity, convergence behavior, and generalization performance.

2.4.3. BO framework

BO models the objective function $f(\boldsymbol{\theta})$ as a Gaussian process (GP), expressed as

$$f(\boldsymbol{\theta}) \sim GP\left(\mu(\boldsymbol{\theta}), k(\boldsymbol{\theta}, \boldsymbol{\theta}')\right) \qquad (16)$$

where $\mu(\boldsymbol{\theta})$ and $k(\boldsymbol{\theta}, \boldsymbol{\theta}')$ denote the mean and covariance functions, respectively. Based on previously evaluated hyperparameter configurations, the GP surrogate provides a probabilistic estimate of the objective function, including both the expected value and associated uncertainty.

To guide the search process, the expected improvement (EI) acquisition function is employed, defined as

$$EI(\boldsymbol{\theta}) = \mathrm{E}[\max(f^* - f(\boldsymbol{\theta}), 0)] \qquad (17)$$

where f^* represents the best observed objective value. The next hyperparameter vector is determined by maximizing the acquisition function:

$$\boldsymbol{\theta}_{n+1} = \arg\max_{\boldsymbol{\theta}\in\Omega} EI(\boldsymbol{\theta}). \qquad (18)$$

This iterative process continues until a predefined stopping criterion, such as the maximum number of evaluations, is reached.

3. CASE STUDY

The study focuses on the turbine no. 1 selected from a six-unit wind farm situated along the south-central coastline of Vietnam. The turbine features a hub height of 114 m and a rotor diameter of 132 m. Operational measurements and power output data were acquired from the SCADA system at 10-minute intervals over the period from June 01 to June 30, 2025. The dataset consists of a total of 4,320 observations. The operational variables acquired from the SCADA system include ambient temperature, generator speed, nacelle position, pitch angle, rotor speed, wind speed, bearing temperature, winding temperature, hydraulic group pressure, and wind power output. However, the bearing and winding temperature variables are excluded from further analysis because they primarily reflect the thermal condition of mechanical and electrical components rather than directly influencing the aerodynamic power conversion process. Moreover, these temperature measurements are strongly correlated with operating load and ambient conditions, which may introduce redundancy and multicollinearity without providing additional predictive information for short-term wind power forecasting.

Table 1. Significant inputs using stepwise regression.

Input	Selected by stepwise	p-value
Ambient temp	true	5.3482e-05
Generator speed	true	1.8914e-08
Nacelle position	true	0.0036724
Pitch angle	true	4.3119e-156
Rotor speed	true	2.1991e-06
Wind speed	true	0
Hidraulic group pressure	false	NaN

Stepwise regression is applied to the SCADA dataset to systematically identify the operating parameters that are most relevant to wind power output. All candidate variables are initially considered as potential predictors, and a bidirectional stepwise procedure is employed, in which variables are iteratively added to or removed from the regression model based on their statistical significance. At each iteration, the inclusion or exclusion of a variable is determined using the corresponding p-value obtained from the F-test, with a predefined significance threshold. This process continues until no additional variables meet the criteria for entry or removal, resulting in a parsimonious regression model. The results are shown in Table 1. The selected parameters are demonstrated in Figure 1. Based on the results, the selected variables including ambient temperature, generator speed, nacelle position, pitch angle, rotor speed, wind speed, and wind power output are subsequently used as inputs for the ARIMAX and LSTM forecasting models.

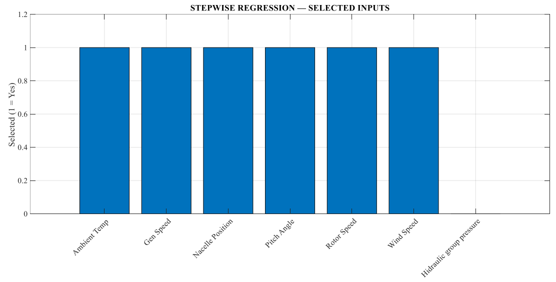

Figure 2. Selected parameters using stepwise regression.

The ambient temperature, generator speed, nacelle position, pitch angle, rotor speed, and wind speed are illustrated in Figures 3, 4, 5, 6, 7, and 8, respectively.

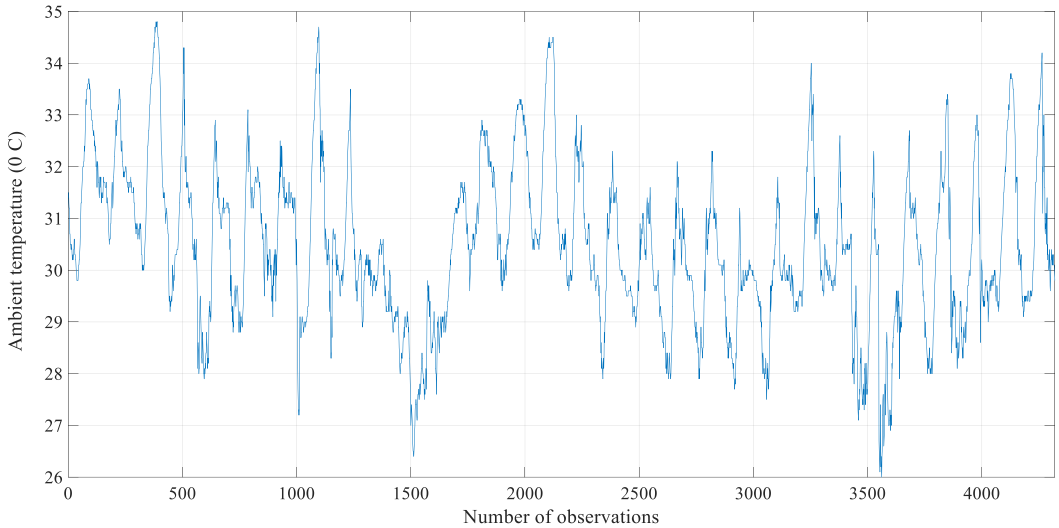

Figure 3. Ambient temperature.

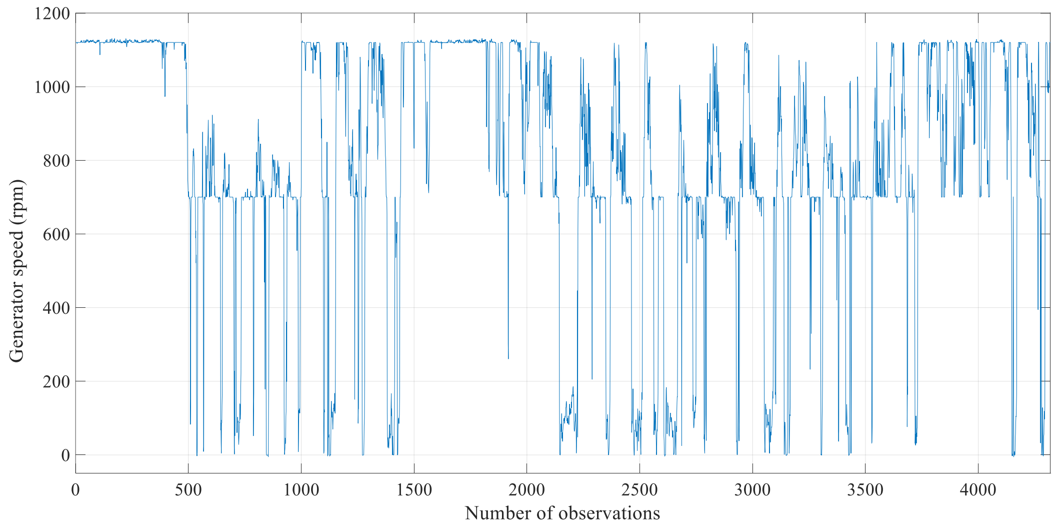

Figure 4. Generator speed.

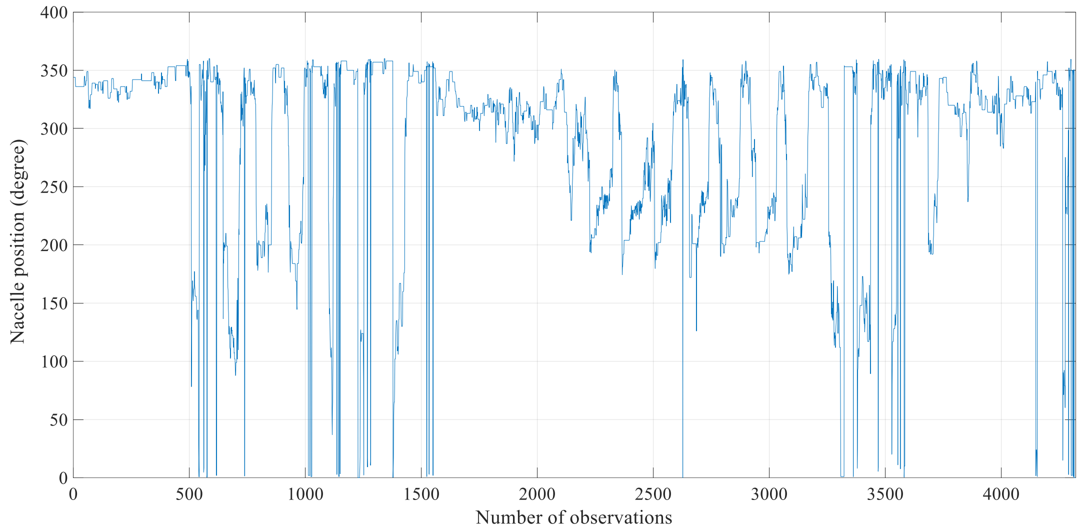

Figure 5. Nacelle position.

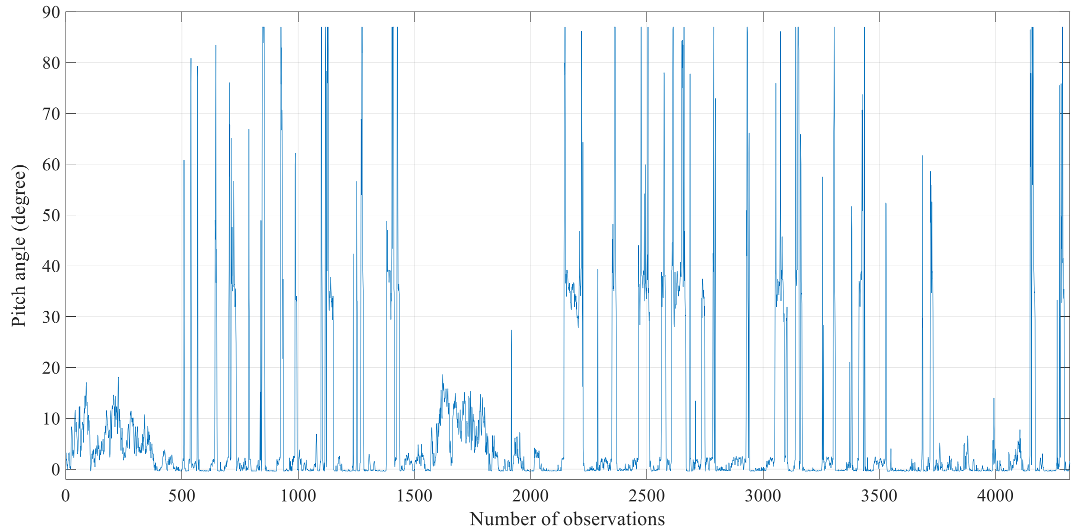

Figure 6. Pitch angle.

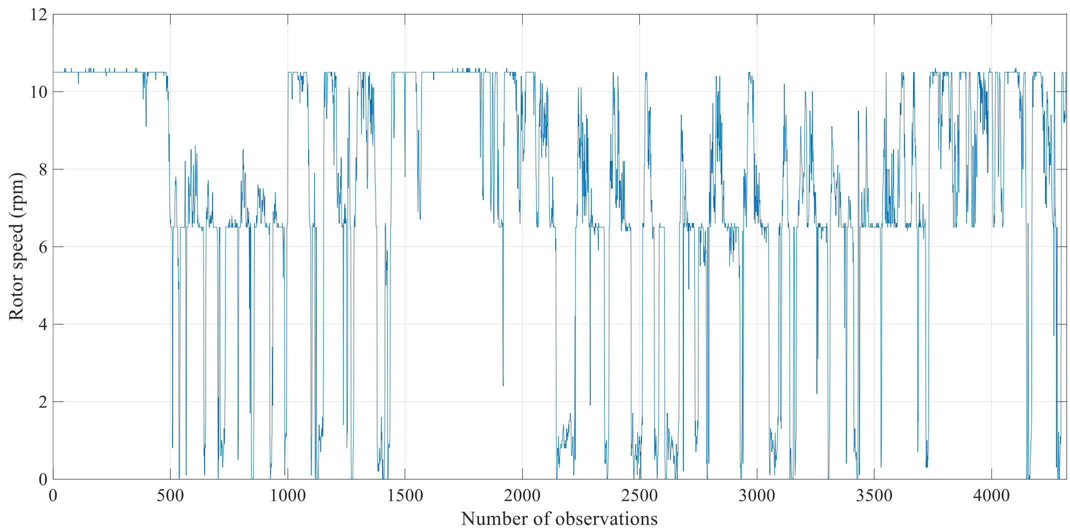

Figure 7. Rotor speed.

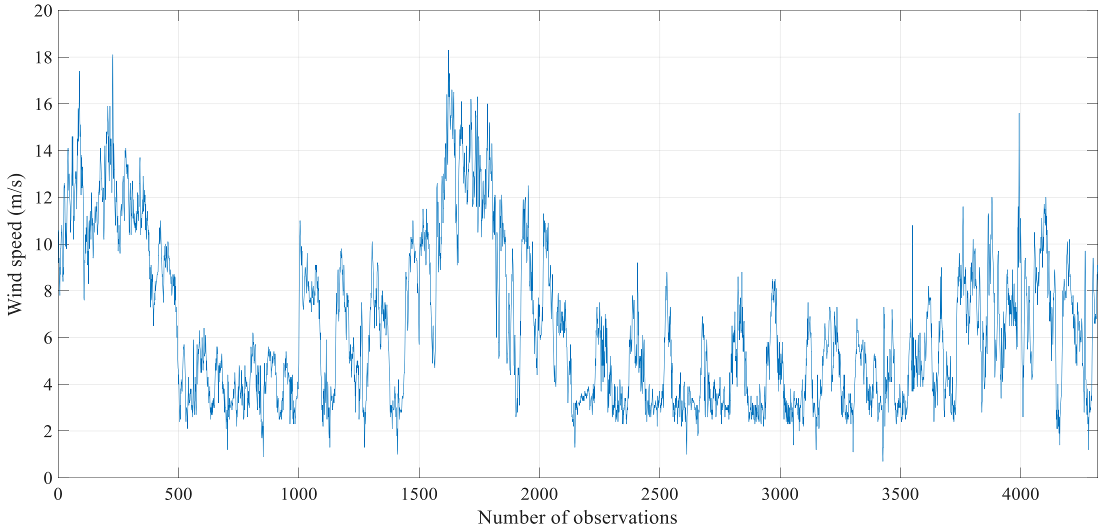

Figure 8. Wind speed.

To establish a statistical benchmark for wind power forecasting, an ARIMAX model is developed using the significant operating parameters identified through the stepwise regression analysis. The model order is determined by evaluating multiple competing ARIMAX configurations and comparing them using the BIC. Based on this criterion, the optimal model structure is identified as ARIMAX(0, 0, 3), which yields the minimum BIC value of −6167.29, indicating a favorable trade-off between model goodness-of-fit and model complexity. Using this configuration, the forecasting performance of the ARIMAX model is subsequently evaluated.

In this study, a LSTM network is employed to model the nonlinear temporal dependencies inherent in short-term wind power time-series data. The LSTM model is constructed using historical wind power observations together with the significant operating parameters selected through stepwise regression. A sliding window approach is adopted to generate input–output pairs, where the look-back window determines the number of previous time steps used for prediction.

To further enhance forecasting accuracy and robustness, BO is applied to automatically tune the key hyperparameters of the LSTM model. The optimization process aims to minimize a composite objective function, defined in Equation (13), which jointly considers absolute and relative forecasting errors through a weighted combination of MAE and WAPE. By varying the weighting coefficient α, the trade-off between these two error metrics is systematically investigated.

The LSTM model using BO is coded and trained in Matlab software. The settings of LSTM are:

- ✓ Number of hidden: [10, 300]
- ✓ Learning rate: [1e-4, 5e-3] (log scale)
- ✓ Number of epochs: [10, 500]
- ✓ Dropout rate: [0, 0.5]
- ✓ Look-back window length: [5, 30]

The settings of BO (bayesopt) are:

- ✓ MaxObjectiveEvaluations: 20
- ✓ IsObjectiveDeterministic: true
- ✓ UseParallel: false

The optimization results indicate that the best overall performance is achieved when $\alpha = 0.6$, which provides a balanced emphasis on both error criteria. Under this setting, BO converges to the following optimal LSTM hyperparameter configuration:

- ✓ Number of hidden units (N_h): 193
- ✓ Learning rate (η): 0.00048928
- ✓ Number of epochs (N_e): 467
- ✓ Dropout rate (d): 0.0094544
- ✓ Look-back window length (L): 27

Using this optimized configuration, the forecasting performance of the BO–based LSTM model is evaluated. Compared with other values of α, the results obtained at $\alpha = 0.6$ demonstrate a superior balance between absolute and relative error minimization, thereby confirming the effectiveness of the proposed composite objective function. The performance evaluation results of all forecasting models considered are summarized in Table 2.

Table 2. Evaluation criteria.

Model	MAE	WAPE (%)
ARIMAX	108.8	9.282
BO–based LSTM	83.81	7.20

As shown in Table 2, the BO–based LSTM model significantly outperforms the ARIMAX benchmark in terms of both MAE and WAPE. Specifically, the LSTM model achieves a reduction of approximately 22.9% in MAE and 22.4% in WAPE compared with the ARIMAX model, indicating its superior capability in capturing nonlinear temporal patterns and complex dependencies in wind power generation data. Figures 9 and 10 illustrate comparisons between the actual wind power generation and the forecasting results produced by the ARIMAX and BO–based LSTM models, respectively. The actual wind power values are represented by solid blue lines, while the corresponding predicted values are shown as dashed red lines. The visual comparison further confirms that the BO–based LSTM model more closely tracks the variations in wind power output, particularly during rapid fluctuations.

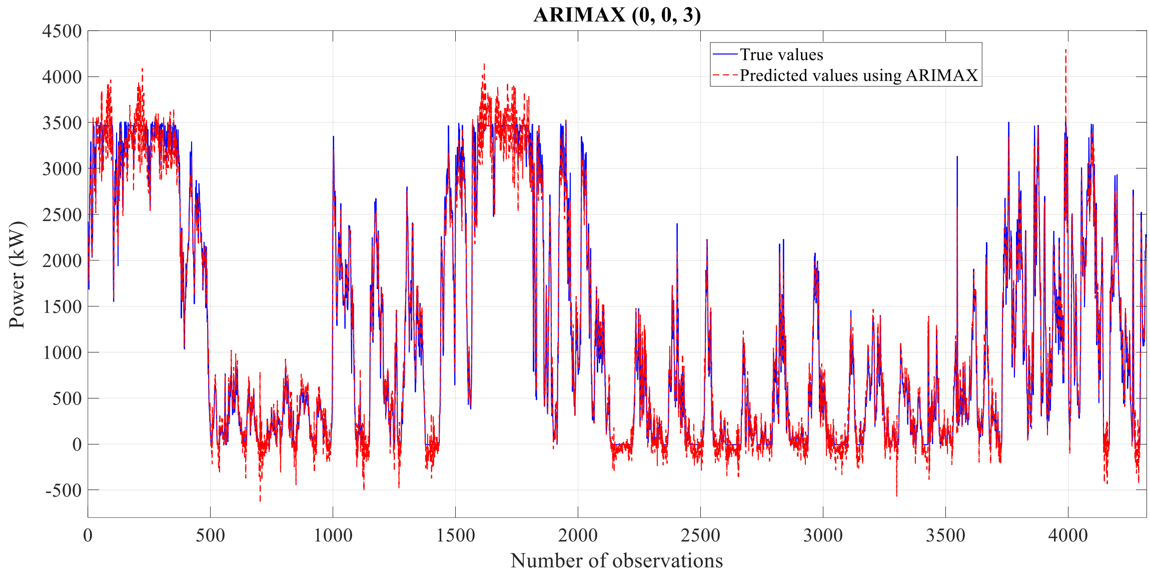

Figure 9. The forecasted results of wind power generation using ARIMAX model.

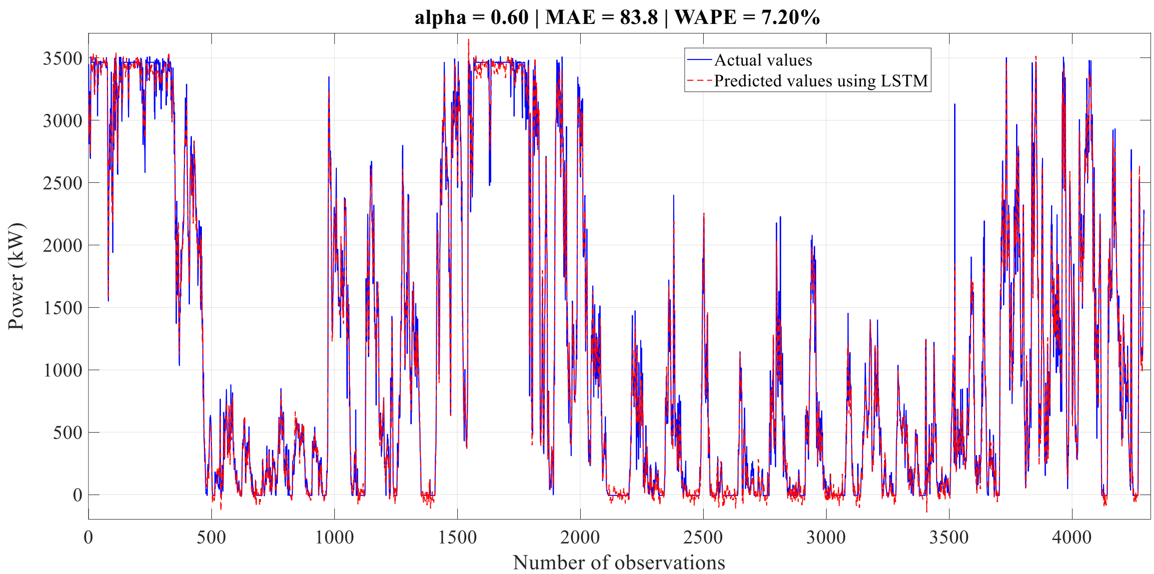

Figure 10. The forecasted results of wind power generation using BO-based LSTM model.

3. CONCLUSION

This paper proposed an effective short-term wind power forecasting framework that leverages high-resolution SCADA data from a practical wind turbine. A stepwise regression method was first applied to identify the most significant operating parameters influencing wind power output, thereby reducing input dimensionality and improving model interpretability. An ARIMAX model incorporating the selected parameters was developed as a statistical benchmark, while an LSTM-based model was employed to capture nonlinear and temporal characteristics of wind power generation. To further enhance forecasting accuracy, BO was integrated to systematically tune the LSTM hyperparameters using a composite objective function combining MAE and WAPE. Results from a real-world case study demonstrate that the BO–based LSTM model significantly outperforms the ARIMAX benchmark in terms of both absolute and relative error metrics. The proposed approach shows superior capability in modeling complex wind power dynamics and tracking rapid output fluctuations. Overall, the findings confirm the effectiveness of combining statistically selected SCADA parameters with deep learning and Bayesian hyperparameter optimization for improving wind power forecasting accuracy. Future work will extend the proposed framework to wind farm–level forecasting and probabilistic modeling to better support power system operation under high wind power penetration. Moreover, future work will adopt a strict chronological training–validation–test splitting strategy using extended multi-season datasets to further improve robustness evaluation.

REFERENCES

1. D. D. Tung, T. H. Le. A statistical analysis of short-term wind power forecasting error distribution, *International Journal of Applied Engineering Research*, **2017**, 12(10), 2306-2311.
2. X. Zheng, T. Jin. A reliable method of wind power fluctuation smoothing strategy based on multidimensional non-linear exponential smoothing short-term forecasting, *IET Renewable Power Generation*, **2022**, 16(16), 3573-3586.
3. A. Kusiak, Z. Zhang. Short-horizon prediction of wind power: A data-driven approach, *IEEE Transactions on Energy Conversion*, **2010**, 25(4), 1112-1122.
4. A. Anwar, A. N. Mahmood. *Enhanced estimation of Autoregressive wind power prediction model using constriction factor particle swarm optimization*, 2014 9th IEEE Conference on Industrial Electronics and Applications, Hangzhou, China, 2014.
5. J. Wang, Q. Zhou, X. Zhang. *Wind power forecasting based on time series ARMA model*, IOP conference series: Earth and environmental science, 2018, 022015.
6. R. B. Magadum, S. Bilagi, S. Bhandarkar, A. Patil, A. Joshi. *Short-term wind power forecast using time series analysis: Auto-regressive moving-average model (ARMA)*, Recent developments in electrical and electronics engineering: select proceedings of ICRDEEE 2022, Singapore: Springer Nature Singapore, 2023, 319-341.
7. F. A. Eldali, T. M. Hansen, S. Suryanarayanan, E. K. P. Chong. *Employing ARIMA models to improve wind power forecasts: A case study in ERCOT*, 2016 North American Power Symposium (NAPS). IEEE, 2016. 1-6.
8. B. M. Hodge, A. Zeiler, D. Brooks, G. Blau, J. Pekny, G. Reklatis. Improved wind power forecasting with ARIMA models, *Computer Aided Chemical Engineering*, **2011**, 29, 1789-1793.
9. G. E. Box, G. M. Jenkins, G. C. Reinsel, G. M. Ljung. *Time series analysis: forecasting and control (5th Edition)*, John Wiley & Sons, 2015.
10. K. George. *Renewable Energy Forecasting: From Models to Applications*, Woodhead Publishing, 2017.
11. J. Li, H. Cheng, X. Zhou, M. Wang, Z. Wang. *Optimization of multi-layer perceptron for wind power generation prediction based on improved grey wolf algorithm*, International Symposium on New Energy and Electrical Technology. Singapore: Springer Nature Singapore, 2023, 547-557.
12. W. Sun, Y. Wang. Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network, *Energy conversion and Management*, **2018**, 157, 1-12.
13. J. Huang, R. Qin. Elman neural network considering dynamic time delay estimation for short-term forecasting of offshore wind power, *Applied Energy*, **2024**, 358, 122671.
14. Z. Xiao, F. Tang, M. Wang. Wind power short-term forecasting method based on LSTM and multiple error correction, *Sustainability*, **2023**, 15(4), 3798.
15. Z. Niu, Z. Yu, W. Tang, Q. Wu, M. Reformat. Wind power forecasting using attention-based gated recurrent unit network, *Energy*, **2020**, 196, 117081.
16. L. Jovanovic, K. Kumpf, N. Bacanin, A. Milos, M. Joseph, S. Hothefa, Z. Miodrag. *Decomposition Aided Bidirectional Long-Short-Term Memory Optimized by Hybrid Metaheuristic Applied for Wind Power*

Forecasting, Computational Sciences and Sustainable Technologies: Springer Nature Switzerland, 2024, 30-42.

17. H. Wang, Z. Lei, Y. Liu, J. Peng, J. Liu. Echo state network based ensemble approach for wind power forecasting, *Energy Conversion and Management*, **2019**, 201, 112188.
18. J. Zeng, W. Qiao. *Support vector machine-based short-term wind power forecasting*, 2011 IEEE/PES power systems conference and exposition, IEEE, 2011. 1-8.
19. S. Park, S. Jung, J. Lee, J. Hur. A short-term forecasting of wind power outputs based on gradient boosting regression tree algorithms, *Energies*, **2023**, 16(3), 1132.
20. D. Kumar, R. Abhinav, N. Pindoriya. *An ensemble model for short-term wind power forecasting using deep learning and gradient boosting algorithms*, 2020 21st national power systems conference (NPSC), IEEE, 2020, 1-6.
21. E. Dokur, M. Kurban, S. Ceyhan. *Hybrid model for short term wind speed forecasting using empirical mode decomposition and artificial neural network*, 9th International Conference on Electrical and Electronics Engineering (ELECO), Bursa, Turkey, 2015, 420-423.
22. Y. Zhang, L. Zhang, D. Sun, K. Jin, Y. Gu. Short-Term Wind Power Forecasting Based on VMD and a Hybrid SSA-TCN-BiGRU Network, *Applied Sciences,* **2023**, 13(17), 9888.
23. X. Yuan, Q. Tan, X. Lei, Y. Yuan, X. Wu. Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine, *Energy*, **2017**, 129, 122-137.
24. Y. Jiang, X. Chen, K. Yu, Y. Liao. Short-term wind power forecasting using hybrid method based on enhanced boosting algorithm, *Journal of Modern Power Systems and Clean Energy*, **2017**, 5(1), 126-133.
25. H. H. Aly. *A proposed hybrid deep learning model for wind power forecasting*, 2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS), IEEE, 2024, 1011-1014.
26. G. Li, J. Shi, J. Zhou. Bayesian adaptive combination of short-term wind speed forecasts from neural network models, *Renewable Energy*, **2011**, 36(1), 352-359.
27. T. J. Le. A combined method for wind power generation forecasting, *Archives of Electrical Engineering*, **2021**, 70(4), 991-1009.
28. V. N. Nguyen, X. T. Nguyen, V. T. Thong, T. B. Doan, N. T. Luong, T. H. Le, Recurrent neural network models for wind power forecasting, *Quy Nhon University Journal of Science*, **2025**, 19(5), 69-80.