Sensors and Materials, Vol. 37, No. 7 (2025) 2851–2876
MYU Tokyo

S & M 4091

Improving the Efficiency of Wind Power Forecasting: Novel Deep Learning Model with Directed Focus Attention Mechanism

Zhixin Yu,[1] Huaichang Liu,[1] Kunyu Liu,[1]
Jingguo Ma,[1] Bingxiang Ji,[2,3] and Lingling Li[2,3]*

[1]Hebei Shengde Infrastructure Construction Development Co., Chengde 067070, China
[2]State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300401, China
[3]State Key Laboratory of Electromagnetic Field and Electrical Apparatus Reliability of Hebei Province, Hebei University of Technology, Tianjin 300401, China

(Received January 5, 2025; accepted April 22, 2025)

Keywords: wind power generation, power prediction, evolutionary algorithm, attention mechanism, deep learning

In the context of energy transition and sustainable development, wind power forecasting technology is crucial for enhancing system dispatch flexibility and economic efficiency and maximizing wind energy utilization. While sensor popularity and big data and AI advancements have enhanced wind power forecasting accuracy, wind power's stochastic and intermittent nature still challenges forecasting precision. Therefore, in this study, we propose a novel wind power forecasting model integrating a deep learning network with directed attention mechanisms. The model utilizes data obtained from the five-element meteorological sensor, the FT-WQX5 sensor, for initial input. Principal component analysis is employed for data preprocessing, the directed focused attention mechanism is introduced to enhance focus on key information, and the enhanced dynamic strategy-based pied kingfisher optimizer (EDS-PKO) is utilized for parameter optimization. The model integrates long short-term memory networks to capture temporal features. Results demonstrate that under normal weather conditions, the proposed model achieves root mean square error (*RMSE*) below 0.8, R-square (R^2) above 90%, and mean absolute percentage error (*MAPE*) below 7%. Under adverse conditions, the model optimized with the improved EDS-PKO algorithm shows approximately 10% improvement in R^2 and around 20% reduction in *RMSE* compared with other comparative models, with *MAPE* as low as 4.55%. This research provides a new technological approach for wind power forecasting, contributing to efficient wind energy utilization and stable grid operation.

1. Introduction

In the context of global energy structural transformation and sustainable development strategies, wind energy, as a clean and renewable energy form, has become a crucial pathway for countries to alleviate energy shortages, reduce greenhouse gas emissions, and promote green

*Corresponding author: e-mail: lilingling@hebut.edu.cn
https://doi.org/10.18494/SAM5545

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

and low-carbon economic development.(1) According to the Global Wind Energy Council (GWEC)'s Global Wind Report 2024 released on April 25, 2024, the global wind energy industry achieved a record-breaking annual installed capacity of 117 GW in 2023, marking a 50% increase compared with that in 2022, making it the best year for new wind energy development.(2) Wind power prediction based on meteorological sensor data is one of the key technologies for efficient wind power integration and grid scheduling, and is of inestimable value in improving the flexibility, stability, and economic efficiency of power system operation. With the rapid advancement of technologies such as big data, artificial intelligence, and cloud computing, the accuracy and efficiency of wind power forecasting have significantly improved. However, owing to the inherent randomness and intermittency of wind power itself, wind power forecasting still faces numerous challenges.(3) In wind power prediction research, sensors, materials, and technology are crucial. Meteorological sensors monitor wind speed, direction, temperature, and humidity in real time, providing essential data for prediction models. Sensor accuracy and stability directly impact data quality and model performance. For instance, sensors with higher precision in measuring wind speed and temperature—characterized by smaller error margins, lower noise levels, and enhanced signal-to-noise ratios—provide superior data inputs, enabling more accurate wind power predictions.(4) Therefore, constructing an efficient and accurate wind energy prediction model based on wind-power-related sensor data is a necessary way to promote the effective utilization of clean wind energy resources and realize the sustainable development of the power industry.

Currently, wind power forecasting methods can be classified into three major categories: physical methods, statistical methods, and machine learning. Physical modeling based on numerical weather prediction (NWP) relies on fundamental theories of wind energy conversion and meteorological expertise. It integrates multidimensional information such as wind speed, wind direction, temperature, and terrain, applying physical laws and empirical models to predict future wind power generation potentials.(5) For instance, Liu *et al.*(6) enhanced NWP techniques by combining rank aggregation and probabilistic fluctuation perception to better identify wind variations and improve wind power forecasting accuracy. Salgado *et al.*(7) established a physical model and employed Kalman filtering to predict power generation. However, Lin and Zhang(8) argued that physical models are inadequate in handling complex nonlinear relationships and have limited adaptability in complex terrain and variable meteorological conditions. Additionally, physical methods impose high demands on data quality, constraining their applicability in wind power prediction.

In contrast to physical methods, statistical methods primarily rely on historical data to predict relationships between meteorological data and wind power data, including autoregressive (AR), AR moving average (ARMA), and AR integrated moving average (ARIMA).(9) However, Aasim *et al.*(10) found ARIMA to be less accurate in predicting high-frequency subsequences. To enhance ARIMA's fitting capability, Kavasseri and Seetharaman(11) proposed a fractional ARIMA model by introducing differential parameters to improve upon traditional ARIMA models. Nevertheless, Yu *et al.*(12) contend that statistical models are based on the assumption of linear relationships between meteorological and wind power data, limiting their ability to handle complex nonlinear relationships and thereby failing to accurately predict highly nonlinear and dynamically complex wind power generation data.

Machine learning methods enhance prediction performance by uncovering implicit patterns and trends in historical data, making it suitable for handling complex nonlinear time series, widely applied in short-term wind power forecasting.(13) Common machine learning methods include random forest (RF), support vector machine (SVM), and extreme learning machine (ELM).(14) By leveraging powerful nonlinear modeling capabilities, machine learning achieves more accurate predictions than physical and statistical methods when dealing with nonlinear data.(15) For instance, Ji *et al.*(16) proposed an SVM-based wind power prediction model integrating chaos analysis theory and phase space reconstruction principles to capture complex mappings between multiple input variables and wind power generation. Similarly, Rayi *et al.*(17) introduced a hybrid kernel ELM autoencoder model combining variational mode decomposition and deep learning for high-precision short-term and multistep wind power prediction, optimizing model parameters using a novel sine-cosine wavelet cycle algorithm. Despite surpassing traditional statistical methods in handling nonlinear problems, machine learning models encounter challenges such as sensitivity to noise in data, which can lead to inaccurate feature predictions.(18) Additionally, Fu *et al.*(19) argued that the stochastic nature of parameter configurations in machine learning models not only affects their data fitness but also diminishes their generalization ability, ultimately impacting overall prediction performance.

In contrast to machine learning methods, deep learning excels at extracting crucial feature information from complex nonlinear time series and has garnered significant attention in wind power forecasting.(20) Common deep learning models include recurrent neural networks (RNNs) and convolutional neural networks (CNNs). While RNNs have been extensively applied in short-term wind power prediction research and practice, they suffer from gradient explosion issues. Consequently, long short-term memory (LSTM) networks and gated recurrent units (GRUs) have been proposed.(21) For example, Liu *et al.*(22) constructed a bidirectional feature fusion network within LSTM to effectively extract global and local features from time series data. CNNs, adept at capturing and processing locally correlated features, were employed by Wang *et al.*(23) with an enhanced model to predict short-term wind power generation, achieving notable results. To enhance performance in handling complex sequences, Guo *et al.*(24) introduced bidirectional LSTM (BiLSTM) capable of capturing forward dependences and understanding contextual information in both directions. However, BiLSTM neglects the challenge of capturing long-term dependences in time series, thus failing to accurately predict long-time-scale sequences. To exploit more representative feature information from data, Xiong *et al.*(25) combined CNN and LSTM with attention mechanisms. Nevertheless, while attention mechanisms focus the model on crucial information, they may inadvertently emphasize incorrect details if input data features are insufficient or ambiguous, thereby neglecting key information and leading to decreased prediction performance. In summary, wind power data in practical applications often exhibit high uncertainty, seasonal variations, and random fluctuations. Enhancing model generalization capabilities to ensure stable prediction performance under diverse environmental conditions remains a pressing challenge in wind power forecasting research.

Previous research has made significant advances in the field of wind power forecasting. However, it still faces several challenges. First, although most wind power forecasting models have made progress in handling nonlinear data, their accuracy in predicting highly nonlinear and dynamically changing wind power data is still affected. Second, traditional methods with stochastic model parameters not only affect their adaptability to different datasets but also weaken the model's generalization ability, posing a serious obstacle to the stability of the model in practical applications. Lastly, while attention mechanisms can explore more representative features in the data, they may overlook key information when faced with highly uncertain, strongly seasonal, and randomly fluctuating wind power data, potentially degrading the model's performance when incorporating attention mechanisms.

To address these challenges, we proposed a directed focus attention-enhanced deep learning network aimed at improving prediction accuracy, reducing forecasting errors, and providing more reliable data support for grid dispatch. The main contributions and innovations are as follows.

(1) Constructed a composite model integrating evolutionary algorithms, a directed focus attention mechanism (DFAT), and deep learning to extract key features from highly nonlinear time-series data of wind power captured by the FT-WQX5 sensor.

(2) Proposed the enhanced dynamic strategy-based pied kingfisher optimizer (EDS-PKO), integrating dynamic parameter regulation mechanisms and two-stage search modes.

(3) Innovatively built a dynamic focus mechanism within the attention mechanism (DFAT), introducing the DFAT. By altering attention scores of different elements through directed focus gate units, it ensures focusing on the key information in various environments.

The rest of this paper is organized as follows. In Sect. 2, we elaborate on the principles of principal component analysis (PCA). In Sect. 3, we discuss the basic principles of the EDS-PKO-DFAT-LSTM model. In Sect. 4, we elucidate the short-term wind power forecasting process based on the EDS-PKO-DFAT-LSTM model. In Sect. 5, we demonstrate the superiority and effectiveness of the proposed model and algorithms through case studies. In Sect. 6, we summarize the main findings and contributions and discuss its limitations.

2. Data Preprocessing

The raw meteorological data for wind power generation typically includes multiple variables collected by meteorological sensors, such as wind speed, wind direction, temperature, and humidity. The accuracy and stability of these sensors directly affect the data quality, which in turn impacts the performance of prediction models. The FT-WQX5 sensor is employed to measure the data required for wind power prediction. Moreover, data preprocessing is equally crucial to ensure that the model learns from high-quality data. Data dimensionality reduction is commonly employed to address this issue. This approach not only effectively resolves redundancy among multiple variables but also significantly reduces the time and resources required for data processing and model training. It also avoids interference from irrelevant or weakly correlated features on the model's predictive performance, ensuring that the model focuses on the most critical variables, thereby enhancing prediction accuracy and stability.

Currently, in the field of data dimensionality reduction, various techniques are widely applied, including PCA, linear discriminant analysis, factor analysis, local linear embedding, and nonnegative matrix factorization. Unlike typical data dimensionality reduction methods, PCA can generate highly interpretable new features—principal components—as linear combinations of the original variables, preserving the primary trends in the data. Furthermore, by selecting principal components that maximize variance, PCA effectively highlights the main structure of the data while suppressing noise interference and improving model robustness. Additionally, PCA's computational process is efficient and mature and is suitable for large-scale data processing. We employed PCA to reduce meteorological data dimensionality. The principle of PCA is as follows.

Assuming there is a dataset $\boldsymbol{X}$ with n samples and m features.

$$\boldsymbol{X} = \begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nm} \end{pmatrix}, \tag{1}$$

where $\mathbf{X} \in \mathrm{R}^{n \times m}$, the original variable indices are $x_1, x_2, \ldots, x_m$, the transformed new variable indices are $d_1, d_2, \ldots, d_t$ ($t \leq m$), and p_{nm} represents the dataset of samples.

The new variables are typically expressed as linear combinations of the original variables:

$$\begin{cases} d_1 = p_{11}x_1 + p_{12}x_2 + \cdots + p_{1m}x_m, \\ d_2 = p_{21}x_1 + p_{22}x_2 + \cdots + p_{2m}x_m, \\ \cdots \\ d_t = p_{n1}x_1 + p_{n2}x_2 + \cdots + p_{nm}x_m, \end{cases} \tag{2}$$

where the new variable index d_i is called the principal component of the original variable indices $x_1, x_2, \ldots$, and x_m and d_{ij} are the coefficients of the original indexed variable x_j (j = 1, 2, …, m) on principal component d_i (i = 1, 2, ..., n). The specific steps of PCA are described as follows.

Step 1: Calculate the correlation matrix $\boldsymbol{R} = (r_{jk})_{p \times p}$.

$$r_{jk} = \frac{\sum_{i=1}^{n}(x_{ij} - \overline{x}_j)(x_{ik} - \overline{x}_k)}{\sqrt{\sum_{i=1}^{n}(x_{ij} - \overline{x}_j)^2}\sqrt{\sum_{i=1}^{n}(x_{ik} - \overline{x}_k)^2}} \tag{3}$$

Assume there is a dataset that forms a matrix of j rows and k columns, where $\overline{x_j}$ represents the average of the j row samples, $\overline{x_k}$ represents the average of the k column samples and r_{jk} is the correlation coefficient between the original variables x_j and x_k.

Step 2: Solve the characteristic equation $|\boldsymbol{R} - \lambda\boldsymbol{E}| = 0$ (where $\boldsymbol{E}$ is the identity matrix) to obtain the initial values and arrange them in descending order as $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$.

Step 3: Calculate the eigenvector $\vec{e_i}$ corresponding to the eigenvalue λ_i:

$$R\vec{e_i} = \lambda_p \vec{e_i}, \tag{4}$$

$$(\vec{e_i})^T \vec{e_i} = 1 \;\; (i = 1, 2, \ldots, p). \tag{5}$$

In PCA, the selection of principal components is based on the contribution rate ACR:

$$ACR = \left(\frac{\sum_{j=1}^{m} \lambda_i}{\sum_{k=1}^{p} \lambda_k} \right) \times 100\%, \tag{6}$$

where the value of m ($m \leq p$) corresponding to $ACR \geq 80\%$ is recognized as the number of principal components.

PCA reduces data dimensionality while retaining maximum information, enhancing the efficiency and accuracy of predictive models, and improving the sensitivity and generalization ability of the model to key variables. The static information of the wind farm and the preprocessing process of the wind power data set based on PCA are shown in Fig. 1.

3. Formulation of Short-term Wind Power Forecasting Model Based on EDS-PKO-LSTM-DFAT

3.1 LSTM neural network

The LSTM neural network is an improved type of RNN that efficiently handles time series data. Compared with traditional RNNs, LSTM introduces the concepts of forget gate, input gate, output gate, and cell state based on "LSTM". It continuously updates the "long-term memory", that is, the cell state C_t, through the forget gate and input gate. The output gate integrates information from the current time step and the "short-term memory" from the previous time step, that is, the hidden state h_{t-1}, to compute the current time step's LSTM network output h_t. The framework structure of LSTM is shown in Fig. 2.

As shown in Fig. 2, the special structure of LSTM addresses the long-term dependence issue of RNN networks, overcomes the gradient explosion problem in RNNs, and enhances its performance in handling longer time series data. It has been widely applied by many scholars in various fields such as language modeling and text generation, financial investment analysis, and time series forecasting. The principles of LSTM are as follows:

$$F_t = sigmoid(A_f x_t + B_f h_{t-1} + b_f), \tag{7}$$

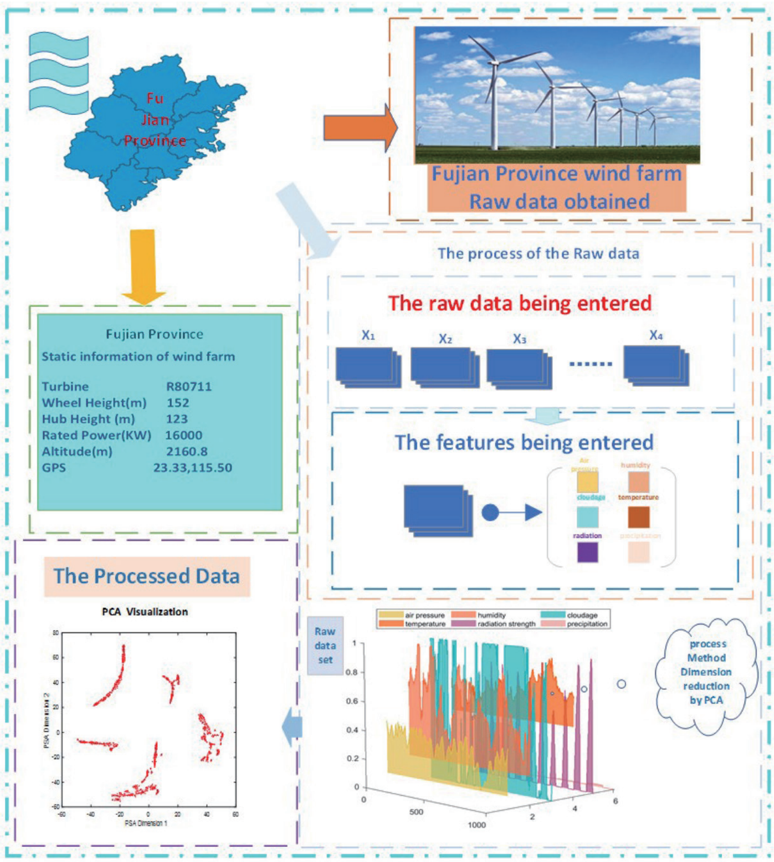

Fig. 1. (Color online) Static information of wind farm and preprocessing process of wind power data set based on PCA.

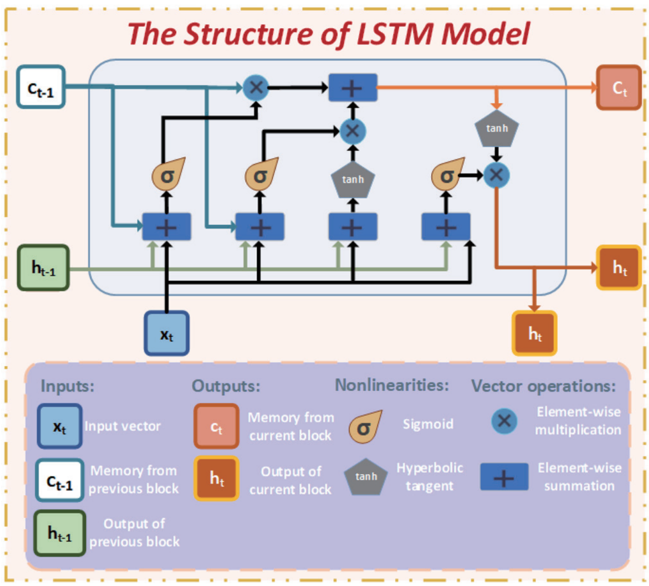

Fig. 2. (Color online) Framework structure of LSTM.

$$I_t = sigmoid(A_i x_t + B_i h_{t-1} + b_i)\,, \tag{8}$$

$$C_t = F_t \times C_{t-1} + I_t \times \tanh(A_c x_t + B_c h_{t-1} + b_c)\,, \tag{9}$$

$$O_t = sigmoid(A_o x_t + B_o h_{t-1} + b_o)\,, \tag{10}$$

$$h_t = O_t \tanh(C_t)\,, \tag{11}$$

where F_t, I_t, and O_t represent the forget gate, input gate, and output gate, respectively; x_t denotes the input data of LSTM at time t; h_{t-1} represents the hidden state at time $t-1$; A_f, B_f, A_i, B_i, A_c, B_c, A_o, and B_o are the weight parameters for the corresponding memory state gates; b_i, b_c, and b_0 are the biases for the corresponding memory state gates; and sigmoid and tanh are activation functions.

3.2 DFAT

As an innovative technique in the field of deep learning, the attention mechanism draws inspiration from the dynamic focusing ability of the human brain when processing information. It selectively allocates more attention resources to the most crucial parts amidst complex and diverse data inputs. This mechanism has a wide range of applications in deep learning models, spanning across various domains such as natural language processing, computer vision, and time series analysis, showcasing its remarkable effectiveness in information filtering and feature emphasis. Particularly in tasks involving the analysis of time series data such as wind power prediction, the combination of LSTM and attention mechanism plays a crucial role. LSTM excels at capturing long-term dependences in sequential data. However, when faced with varying levels of information importance across consecutive time steps, relying solely on LSTM may not efficiently utilize all the data points. In such cases, introducing the attention mechanism becomes an important means of optimizing prediction performance.

Specifically, applying the attention mechanism to the output sequence of the LSTM model means that the model can actively "weight" the hidden states at each time step on the basis of specific attention scores, distinguishing which data points are more critical for predicting future wind power. This dynamic weighting strategy allows the model to flexibly adjust the emphasis on different historical moments. Time points that are considered more influential for the prediction target are assigned higher weights, amplifying their significance during the information integration process. By adding the attention mechanism, the LSTM model can predict the future wind power value more accurately, which lays a foundation for the subsequent scheduling. This not only enhances the model's sensitivity to dynamically changing meteorological information but also enables the model to accurately grasp the core factors driving wind power variations, ultimately leading to a significant improvement in prediction accuracy.

The attention mechanism calculates the similarity between the query vector and each element of the input vector, and obtains the attention weights for different elements through normalization. The specific calculation formula is

$$S_i = h_i^T Q, \tag{12}$$

$$softmax(W_i) = \frac{e^{S_i}}{\sum_{i=1}^{N} e^{S_i}}, \tag{13}$$

$$\boldsymbol{Y} = \boldsymbol{W} \cdot \boldsymbol{H}, \tag{14}$$

where h_i represents the ith element of the LSTM network output vector; Q represents the query vector; W_i is the attention weight of the ith element; $\boldsymbol{W}$ is the attention weight vector; $\boldsymbol{H}$ is the LSTM output vector; and $\boldsymbol{Y}$ is the output vector of the attention layer.

Although traditional attention mechanisms have shown excellent performance in many scenarios, especially when dealing with highly similar information, effectively filtering out irrelevant distractions and focusing on key features to enhance model prediction performance, their fixed-weight allocation strategy reveals limitations when facing drastic environmental changes. This limitation is particularly evident in applications, such as wind power prediction, that highly depend on environmental factors, where slight fluctuations and extreme variations have distinct impacts on the prediction results. For example, in a relatively stable wind field environment, traditional attention mechanisms tend to emphasize the processing of highly similar data features, which helps filter out occasional noise interference and ensure prediction stability. However, when encountering extreme conditions such as severe weather, the significant fluctuations in environmental factors carry a wealth of potential information about wind power variations. In such cases, continuing to rely on the existing mechanism and overemphasizing similarity matching may lead to missing these exceptional data points that contain crucial information, resulting in inaccurate predictions.

To address this challenge, we innovatively incorporate the concept of dynamic focusing into the attention mechanism framework and design a novel DFAT. The core of this mechanism lies in its ability to determine the volatility of the current environment on the basis of the variance of similarity scores of input data. Variance, as a statistical measure of data dispersion, can sensitively reflect the diversity and intensity of information fluctuations. When the variance exceeds a predefined threshold, it indicates that the environment has entered a highly dynamic phase, such as encountering severe weather. Based on this, the directed focusing gate control unit plays a crucial role by dynamically adjusting the attention weights of each data element according to the judgment result. During periods of environmental calmness, the original weighting strategy is maintained to ensure prediction robustness. However, during periods of environmental turbulence, the weight distribution is actively adjusted to give higher attention to abrupt changes and high-variance-indicated abnormal data, thereby capturing the crucial information changes for wind power prediction under extreme conditions.

Through this dynamic adjustment strategy, the directed focusing attention mechanism not only overcomes the limitations of traditional attention mechanisms in adapting to complex environments but also considerably enhances the model's sensitivity to key information. Whether under calm or severe weather conditions, it ensures that the model focuses on the most relevant input features, effectively improving the accuracy and reliability of wind power prediction. This provides advanced technical support for addressing the ever-changing challenges in practical wind energy management.

The mathematical model of the directed focus gate control unit is

$$D_i = sigmoid(A_i(1-W_i)+b_i) \quad \text{if} \quad \mathrm{var}(W_i > T_{cir}), \tag{15}$$

where D_i represents the updated directed focusing attention weight; A_i and b_i are the weight parameters and biases of the directed focusing gate, respectively; and T_{cir} is the threshold for determining severe weather conditions, which is determined on the basis of historical wind power generation data (when the variance of attention scores exceeds this threshold, it is considered as severe weather). The structure of LSTM-DFAT is presented in Fig. 3.

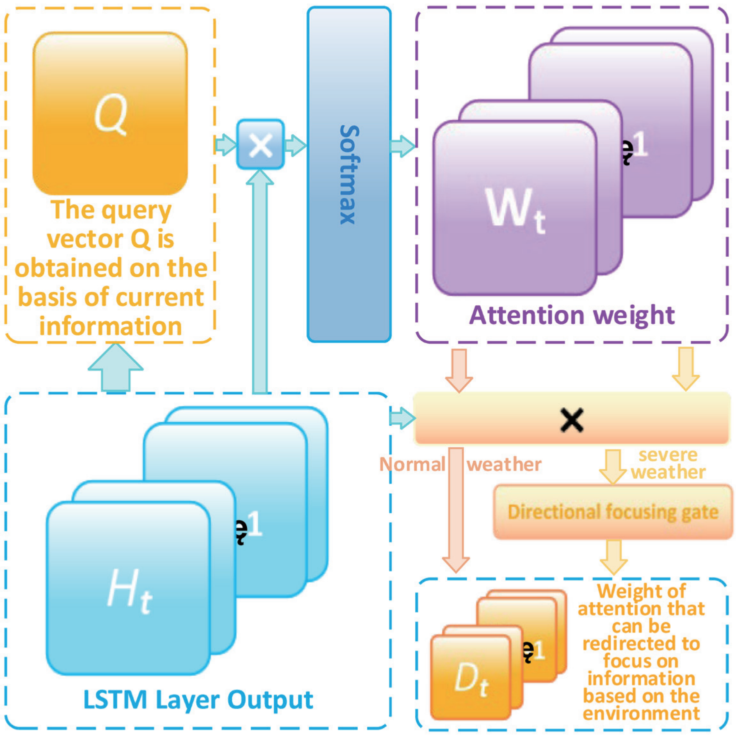

Fig. 3. (Color online) Structure of LSTM-DFAT.

The DFAT enhances the model's prediction performance in complex and changing environments through its dynamic adjustment characteristics. It not only improves the accuracy and reliability of predictions but also enhances the flexibility and practicality of the model. This provides a more advanced and reliable tool for wind power energy management and scheduling.

3.3 Kingfisher class optimization algorithm

The pied kingfisher optimizer (PKO) is an emerging bio-inspired optimization algorithm inspired by the hunting behavior of the pied kingfisher. The algorithm consists of three main stages: perching/hovering, diving hunting, and symbiosis. By simulating the foraging strategies of the pied kingfisher in nature, the PKO algorithm demonstrates excellent performance in solving complex optimization problems. The following are the basic strategies and steps of the PKO algorithm.

3.3.1 Perching/hovering stage

In this stage, the pied kingfisher observes its prey by perching and hovering above the water surface. In the algorithm, the initial population is randomly generated and distributed in the search space. The main objective of this stage is to explore the search space and increase the diversity of the population to ensure the possibility of finding the global optimum solution.

$$X_i(t+1) = X_i(t) + \alpha * T \times \left(X_j(t) - X_i(t)\right),\ \alpha \in 2 * randn(1, \dim) \tag{16}$$

Here, N is the total population size; i and j are natural numbers between 1 and N that are not equal to each other; *randn* is a random number that follows a normal distribution; and Dim represents the dimension of the problem under consideration.

The calculation of the T parameter in the perching strategy is as follows:

$$T = \left(\exp(1) - \exp\left(\frac{t-1}{Max\_iter}\right)^{\frac{1}{BF}} \right) * \cos(Crest), \tag{17}$$

$$Crest = 2 * pi * rand, \tag{18}$$

where *Max_iter* is the maximum number of iterations; and *rand* is a random value between 0 and 1.

The calculation of the parameter T in the hovering strategy is as follows:

$$T = beating * \left(\frac{t^{\frac{1}{BF}}}{Max\_iter^{\frac{1}{BF}}} \right), \tag{19}$$

$$beating = rand*\left(\frac{PKO\_Fitness(i)}{PKO\_Fitness(j)}\right), \tag{20}$$

where *PKO_Fitness* represents the fitness value of the pied kingfisher, and *BF* (*BF*=8) is the bouncing factor.

3.3.2 Diving hunting stage

When the prey is detected, the pied kingfisher quickly dives to hunt. In the PKO algorithm, this stage corresponds to the deep exploitation of potential good solutions. By adjusting the individuals' positions, they are brought closer to the optimal solution. The specific steps are as follows.

(1) Select excellent individuals: On the basis of their fitness values, select a subset of excellent individuals as diving targets.

(2) Update positions: using the chosen diving strategy, update the individuals' positions to make them closer to the local optimum. Common methods include linear weighting strategies or nonlinear adjustment strategies to ensure the convergence of the algorithm. For each individual, the position update formula is

$$X_i^{t+1} = X_i^t + r \cdot \left(X_{best}^t - X_i^t\right) + s \cdot randn(0,1), \tag{21}$$

where X_{best}^t is the best position in the current population; r and s are control parameters; and *randn*(0, 1) is a random number that follows a normal distribution with a mean of 0 and a standard deviation of 1.

3.4 EDS-PKO

Owing to the inherent structure of the algorithm, the original PKO algorithm is prone to falling into local optima and faces challenges in balancing global exploration with local exploitation. To address this, we introduced an enhanced dynamic strategy, including dynamic parameter adjustments, the flight factor, and the two-stage update strategy.

(1) Dynamic Adjustment of Beating Factor (BF) and Flight factor

In nature, pied kingfishers adjust their hunting strategies on the basis of different environments and prey. In PKO, by dynamically adjusting the BF to gradually decrease with the number of iterations, it enhances exploratory capabilities in the early stages and exploitation capabilities in the later stages.

$$BF = 8 \cdot \left(1 - \frac{t}{Max\_iter}\right) \tag{22}$$

BF is kept at a high value in the early stages to promote extensive exploration of the search space, gradually decreasing as the iterations progress to focus the algorithm on local search.

When pied kingfishers hunt while flying, they adjust their flight trajectory by changing their speed and direction. To simulate this process, a flight factor (w) is introduced, and a dynamic adjustment strategy is adopted to ensure that the algorithm has strong global exploration capabilities in the early stages and stronger local exploitation capabilities in the later stages.

$$w = 0.9 - \left(0.5 \cdot \left(\frac{t}{Max\_iter} \right) \right) \tag{23}$$

(2) Two-stage Update Strategy

When pied kingfishers hunt, they first engage in extensive exploration to locate prey and then focus their efforts on hunting. To enhance the stability and diversity of the algorithm, we introduced a two-stage update strategy, employing different update methods for the exploration and exploitation stages.

During the exploration stage, individuals primarily conduct extensive searches to discover more potential optimal solutions:

$$\begin{cases} X_i^{t+1} = w \cdot X_i^t + \alpha \cdot \left(\dfrac{rand \cdot f\left(X_j\right)}{f\left(X_i\right)} \text{-} \left(\dfrac{t}{Max\_iter} \right)^{1/BF} \right) \cdot ..., \\ \left(X_j^t - X_i^t \right) + 1.5 \cdot rand \cdot \left(X_{best}^t - X_i^t \right), \\ X_i^{t+1} = w \cdot X_i^t + \alpha \cdot \left(\exp(1) - \exp\left(\dfrac{t-1}{Max\_iter} \right)^{1/BF} \right) \cdot ..., \\ \cos(\theta) \cdot \left(X_j^t - X_i^t \right) + 1.5 \cdot rand \cdot \left(X_{best}^t - X_i^t \right). \end{cases} \tag{24}$$

In the exploitation phase, individuals mainly conduct local searches to refine the accuracy of the current solution:

$$X_i^{t+1} = X_i^t + HuntingAbility \cdot o \cdot \alpha \cdot \left(b - X_{best}^t \right), \tag{25}$$

$$HuntingAbility = \frac{rand \cdot f\left(X_i\right)}{f\left(X_{best}\right)}, \tag{26}$$

$$b = X_i^t + o^2 \cdot randn \cdot X_{best}^t, \tag{27}$$

$$X_i^{t+1} = \begin{cases} X_{rand}^t + o \cdot \alpha \cdot \left| X_i^t - X_{rand}^t \right|, & \text{if } rand > (1 - PE) \\ X_i^t. & \text{else} \end{cases} \tag{28}$$

4. EDS-PKO-LSTM-DFAT-based Short Wind Power Forecasting Process

The process of short-term wind power forecasting based on the EDS-PKO-LSTM-DFAT model is as follows.

(1) The required measurement information for prediction was obtained using the FT-WQX5 sensor. Subsequently, to extract key feature information from the wind power time series, the dimensionality of the original meteorological data was reduced through the application of PCA in the data dimensionality reduction process.

(2) In the model parameter initialization stage, the EDS-PKO-LSTM-DFAT model's parameters, including those of the EDS-PKO algorithm and LSTM-DFAT model, are initialized and shown in Table 1.

(3) Define the fitness function. The degree of fit R^2 between the actual and predicted wind power values is used as the fitness function for the EDS-PKO algorithm.

(4) Optimize model hyperparameters by splitting data into training (first 8 days) and test (ninth day) sets. Use the EDS-PKO algorithm to find the best hyperparameter configuration for the LSTM-DFAT model through training and analysis.

(5) The LSTM-DFAT model, configured with optimal hyperparameters, predicts short-term wind power on the test set to obtain the results.

The process of short-term wind power forecasting based on the EDS-PKO-LSTM-DFAT model is illustrated in Fig. 4.

5. Case Analysis

To comprehensively validate the performance of this model under different weather conditions, especially in adverse weather, we utilized detailed meteorological and wind power generation data from a province in China. Three cases were presented to demonstrate the effectiveness of the model constructed. Since the original dataset did not include explicit weather labels, a classic algorithm in the unsupervised learning domain, K-means clustering, was employed to cluster the meteorological data, and feature analysis was conducted to distinguish between sunny, cloudy, rainy, and adverse weather conditions. The processing results are shown in Fig. 5.

In Case 1, the effectiveness of the EDS-PKO algorithm was validated on the basis of benchmark test functions, and the convergence performance of the algorithm was evaluated

Table 1
Parameter settings.

Parameters		Values
LSTM-DFAT	Input dimension	2
	Output dimension	1
	Predicted quantity of samples	96
	Mini batch size	12
EDS-PKO	Optimizer	Adam
	Loss function	R^2

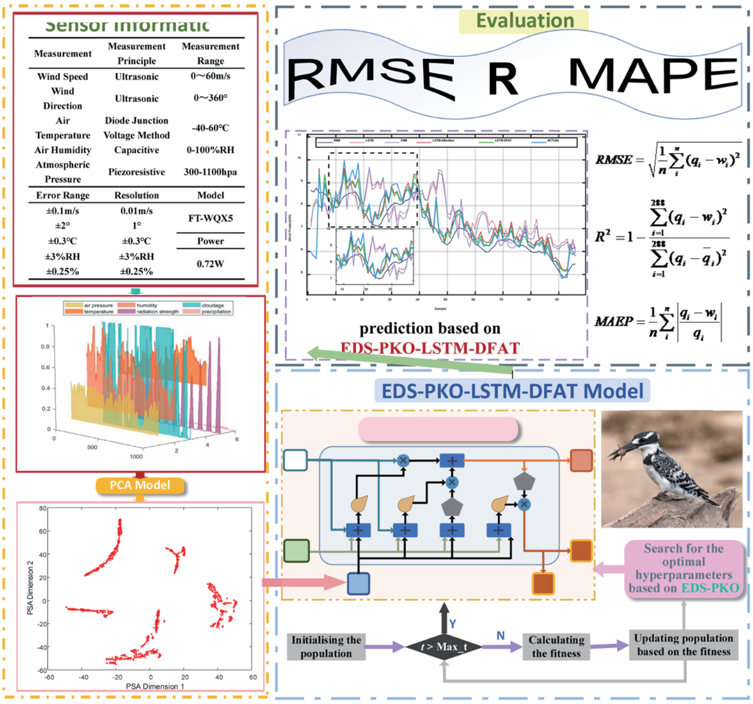

Fig. 4. (Color online) Process of short-term wind power forecasting using the EDS-PKO-LSTM-DFAT model.

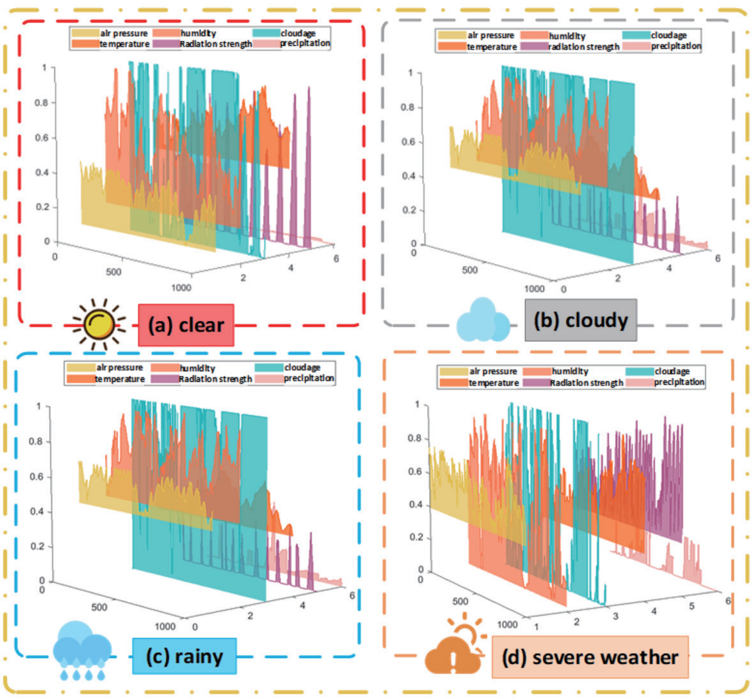

Fig. 5. (Color online) Weather resolution results.

using evaluation metrics. In Case 2, traditional prediction models such as CNN, RNN, LSTM, and LSTM-Attention were used for comparison. The time correlation method was employed for prediction, where data from the past nine days were used as the training set to predict the wind power on the following day under sunny, cloudy, and rainy weather scenarios. In Case 3, the meteorological correlation method was used to predict data under adverse weather conditions for comparison. Owing to the weak temporal correlation of wind power data under adverse weather, using the wind power data from the adverse weather conditions closest to the current time as the training set yielded better results.

5.1 Case 1: Numerical validation and statistical analysis

In existing studies, benchmark test functions are commonly used for the numerical validation of algorithms. Common single-modal and multimodal test functions are typically employed to assess the algorithm's global search performance and its ability to escape local optima. Therefore, we selected multimodal and single-modal test functions to validate the EDS-PKO algorithm's performance in solving different modal problems. Specific information on the selected benchmark test functions is provided in Table 2.

Three widely recognized intelligent optimization methods, namely, the cuckoo search algorithm (CS), grey wolf optimizer (GWO), and particle swarm optimization (PSO), were

Table 2
Benchmark function.

Attribute	Function	Testing dimension	Optimal value	Range
Unimodal	$F_1(k)=\sum_{j=1}^{l}k_j^2$	30	0	[−100, 100]
Unimodal	$F_2(k)=\prod_{j=1}^{l}\lvert k_j\rvert+\sum_{j=1}^{l}\lvert k_j\rvert$	30	0	[−10, 10]
Unimodal	$F_3(k)=\sum_{j=1}^{l}(\sum_{m=1}^{j}k_m)^2$	30	0	[−100, 100]
Multimodal	$F_4(k)=\sum_{j=1}^{l}(k_j^2-10(\cos(2\pi k_j)+10))$	30	0	[−5.12, 5.12]
Multimodal	$F_5(k)=-\exp(\frac{1}{l}\sum_{j=1}^{l}\cos(2\pi k_j))+20+e...$ $-20\exp(-0.2\sqrt{\frac{1}{l}\sum_{j=1}^{l}k_j^2})$	30	0	[−32, 32]
Multimodal	$F_6(k)=\frac{\sum_{j=1}^{l}k_j^2}{4000}-\prod_{j=1}^{l}\cos\left(\frac{k_j}{\sqrt{j}}\right)+1$	30	0	[−600, 600]

selected as comparative algorithms. The parameter settings for each algorithm are detailed in Table 3.

In Table 3, Max_iter refers to the maximum number of iterations for the population, Pop represents the population size, Pa denotes the probability of target discovery in the CS algorithm, BF represents the jump factor in the PSO algorithm, and P represents the initial value of the inertia weight in the EDS-PKO algorithm.

To ensure the accuracy of algorithm performance evaluation, we conducted 30 independent runs of the selected test functions, recording the average value (Avg) and standard deviation (Std) of the experimental results. The numerical verification results of each algorithm are listed in Table 4.

Table 4 shows that, compared with other algorithms, the convergence results of the EDS-PKO algorithm were the most competitive, whether for single-modal or multimodal test functions. With the exception of the standard deviation of test function F1 and the average value of F5, the EDS-PKO algorithm converged to the theoretical optimum of 0, demonstrating excellent convergence performance.

5.2 Case 2: Comparison of different prediction models in normal weather

In Case 2, the wind power prediction under normal weather conditions (sunny, cloudy, rainy) was conducted. CNN, RNN, LSTM, and LSTM-Attention were selected as comparative models to evaluate the superiority of the EDS-PKO-LSTM-DFAT model in handling conventional meteorological situations. All models used the same historical data length, utilizing wind power generation data from the past nine days as input features to predict the wind power output for the next day. The specific implementation steps are as follows.

Table 3
Parameter settings.

Algorithm	Parameters
CS	$Max\_iter = 500$; $Pop = 30$; $Pa = 0.25$
GWO	$Max\_iter = 500$; $Pop = 30$
PSO	$Max\_iter = 500$; $Pop = 30$; $\omega = 0.9$
EDS-PKO	$Max\_iter = 500$; $Pop = 30$; $P = 0.9$

Table 4
Numerical test results.

F	CS		GWO		PSO		EDS-PKO	
	Avg	*Std*	*Avg*	*Std*	*Avg*	*Std*	*Avg*	*Std*
F_1	2.41×10^{-3}	1.26×10^{-3}	6.39×10^{-59}	1.34×10^{-58}	4.46×10^{-5}	5.41×10^{-5}	0	3.03×10^{-72}
F_2	0.33	0.21	9.91×10^{-35}	1.39×10^{-34}	8.25×10^{-4}	1.16×10^{-3}	0	0
F_3	294.36	1.7	1.26×10^{-15}	2.76×10^{-15}	1.24×10^{3}	2.01×10^{3}	0	0
F_4	75.59	10.37	1.44	4.72	49.81	16.80	0	0
F_5	2.54	1.47	1.61×10^{-14}	4.13×10^{-15}	5.84×10^{-2}	0.27	5.66×10^{-16}	0
F_6	6.54×10^{-2}	4.41×10^{-2}	3.83×10^{-4}	2.10×10^{-3}	1.85×10^{-2}	1.93×10^{-2}	0	0

First, each model was trained and tested to ensure that the training and testing sets were divided on the basis of time series to avoid issues related to future information leakage.

Second, root mean square error (*RMSE*), R-square (R^2), and mean absolute percentage error (*MAPE*) were selected as key metrics, to comprehensively assess the model's prediction accuracy.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n}(q_i - w_i)^2} \tag{29}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{288}(q_i - w_i)^2}{\sum_{i=1}^{288}(q_i - \overline{q}_i)^2} \tag{30}$$

$$MAPE = \frac{1}{n}\sum_{i}^{n}\left|\frac{q_i - w_i}{q_i}\right| \tag{31}$$

Here, q_i is the actual value of wind power, w_i is the predicted value of wind power, and $\overline{q}_i$ is the average value of actual wind power. The parameter settings for each prediction model are shown in Table 5.

Table 5
Parameters of each prediction model.

Prediction model	Parameters	Value
CNN	Max Epochs	500
	Initial Learn Rate	0.005
	Mini Batch Size	12
RNN	Max Epochs	500
	Initial Learn Rate	0.005
	Mini Batch Size	12
	Number of RNN layers	2
LSTM	Max Epochs	500
	Initial Learn Rate	0.005
	Mini Batch Size	12
	Number of LSTM layers	2
LSTM-Attention	Max Epochs	500
	Initial Learn Rate	0.005
	Mini Batch Size	12
	Number of LSTM layers	1
	Number of Attention layers	1
LSTM-DFAT	Max Epochs	500
	Initial Learn Rate	0.005
	Mini Batch Size	12
	Number of LSTM layers	1
	Number of PEAN layers	1

For Case 2, the wind power prediction results of each model under three normal weather conditions are shown in Fig. 6.

Figures 6(a)–6(c) depict the prediction results of each model under different weather conditions. Compared with other models, the DFAT-LSTM model's prediction curve aligned most closely with the actual wind power curve under various weather conditions. Figures 6(b), 6(d), and 6(f) show the mean squared error box plot of the prediction results for each model. In

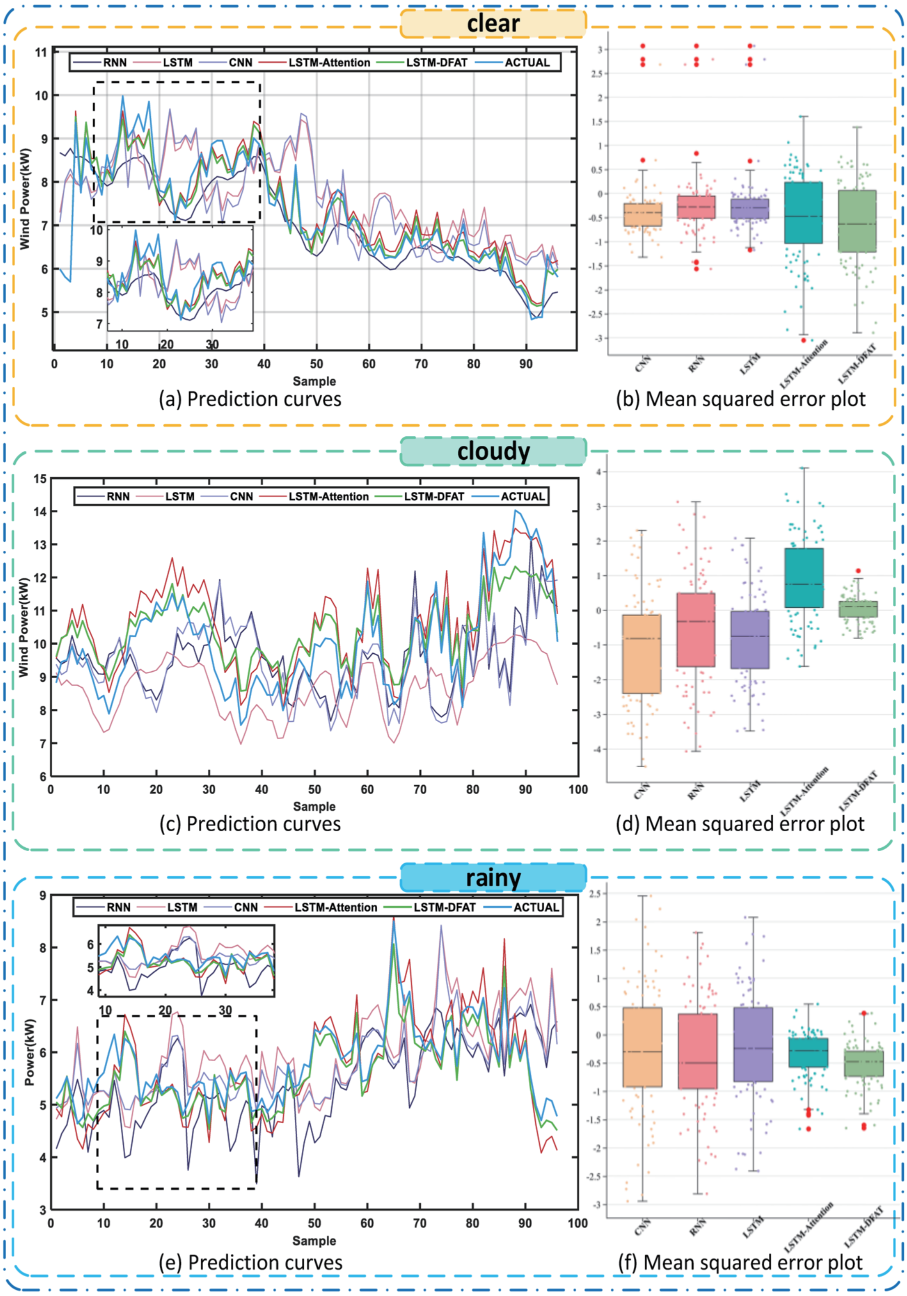

Fig. 6. (Color online) Wind power prediction results for Case 2.

Fig. 6(b), it can be observed that the DFAT-LSTM model did not have significant outlier values in terms of errors. Although the overall mean and variance of prediction errors were relatively large, the prediction performance was still superior to those of other models. Regarding Fig. 6(f), it is evident that while the LSTM-DFAT model had some outliers, these outliers did not differ significantly from the prediction errors of other models under normal circumstances. Additionally, when benchmarked against other traditional models presented in Fig. (6), the LSTM-DFAT model demonstrated notably smaller mean and variance of prediction errors, reflecting its superior prediction performance in this comparative analysis. The LSTM-DFAT model demonstrated broad applicability for wind power under different weather conditions. The summary of prediction error results for each model in Case 2 is presented in Table 6.

Analyzing the data in Table 6 revealed that the LSTM-DFAT model demonstrated higher prediction accuracy across various weather conditions than did traditional models. LSTM-DFAT exhibited outstanding adaptability under the three normal weather conditions, with performance metrics for all predictions surpassing those of traditional models, except for a slightly lower R^2 value under sunny conditions than that of the LSTM-Attention model. Particularly noteworthy was that the *RMSE* was consistently below 0.8, R^2 was above 90%, and *MAPE* was below 7, far superior to those of the RNN, LSTM, and CNN models. These results indicated the LSTM-DFAT model's highly precise capability in predicting wind power.

The results for Case 2 strongly demonstrated the effectiveness and advantages of the LSTM-DFAT model in predicting wind power under normal weather conditions, especially in enhancing prediction accuracy and generalization ability. This provides more reliable decision support for grid dispatch and energy management.

Table 6
Prediction indicators of different models.

Weather Condition	Prediction Models	Assessment criteria		
		R^2 (%)	*MAE*	*MAPE* (%)
Clear	RNN	0.39	1.02	11.34
	LSTM	0.68	0.72	6.63
	CNN	0.40	1.00	11.78
	LSTM-Attention	95.32	0.26	3.02
	LSTM-DFAT	96.48	0.23	2.45
Cloudy	RNN	7.92	1.73	12.92
	LSTM	9.33	1.64	12.60
	CNN	78.93	1.63	13.09
	LSTM-Attention	89.43	0.87	7.21
	LSTM-DFAT	91.37	0.73	6.55
Rain	RNN	5.57	1.03	14.73
	LSTM	1.61	1.01	14.16
	CNN	3.02	0.89	11.93
	LSTM-Attention	91.92	0.37	5.25
	LSTM-DFAT	92.24	0.27	3.60

5.3 Case 3: Comparison of different prediction models in severe weather

In adverse weather conditions, the randomness and volatility of wind power generation are more pronounced, making it beneficial to introduce evolutionary algorithm search to optimize the prediction model's generalization. To further validate the predictive performance of the proposed model under adverse weather conditions and demonstrate the enhanced search capability of the improved EDS-PKO algorithm. We benchmarked the LSTM-DFAT model's predictive performance by subjecting it to hyperparameter optimization via multiple traditional evolutionary algorithms, then rigorously comparing the resultant configurations' forecasting capabilities in Case 3. Table 7 presents the parameter settings of each algorithm.

The wind power prediction results for Case 3 are presented in Fig. 7. In Fig. 7(a), it is evident that under the optimal hyperparameters found by EDS search, the LSTM-DFAT model's prediction closely aligned with the actual wind power curve, demonstrating the precision and effectiveness of the EDS-PKO-LSTM-DFAT model in predicting wind power under adverse weather conditions. The correlations in Fig. 7(b) showed that the scatter plot of the EDS-PKO-LSTM-DFAT model was closest to the 1:1 line, indicating the model's superior correlation between predicted and actual values. Figure 7(c) shows a normalized radar chart of prediction evaluation metrics for each evolutionary algorithm. From this radar chart, it is visually apparent that under the EDS-PKO algorithm search, the LSTM-DFAT model had the highest R^2 value, indicating the optimal correlation between predicted and actual values, while the *RMSE* and *MAPE* metrics were relatively minimal, signifying the smallest prediction errors for the EDS-PKO-LSTM-DFAT model. Table 8 provides the evaluation metrics for Case 3.

Table 8 shows that compared with other models, the proposed EDS-PKO-LSTM-DFAT prediction model excels in all metrics. Specifically, compared with the three baseline models, the R^2 metric showed respective improvements of 16.04, 18.59, and 6.96%, while the *RMSE* metric exhibited respective increases of 42.59, 46.55, and 16.21% for each comparison. The *MAPE* metric for the EDS-PKO-LSTM-DFAT model under adverse weather conditions reached 4.55%, indicating minimal prediction errors when forecasting under adverse weather conditions

Table 7
Search parameter settings of different algorithms for Case 3.

Algorithm	Parameters	Value
PSO	*Max_iter*	20
	Pop	20
	ω	0.9
WOA	*Max_iter*	20
	Pop	20
	e	1
GWO	*Max_iter*	20
	Pop	20
EDS-PKO	*Max_iter*	20
	Pop	20
	Period Threshold	0.8

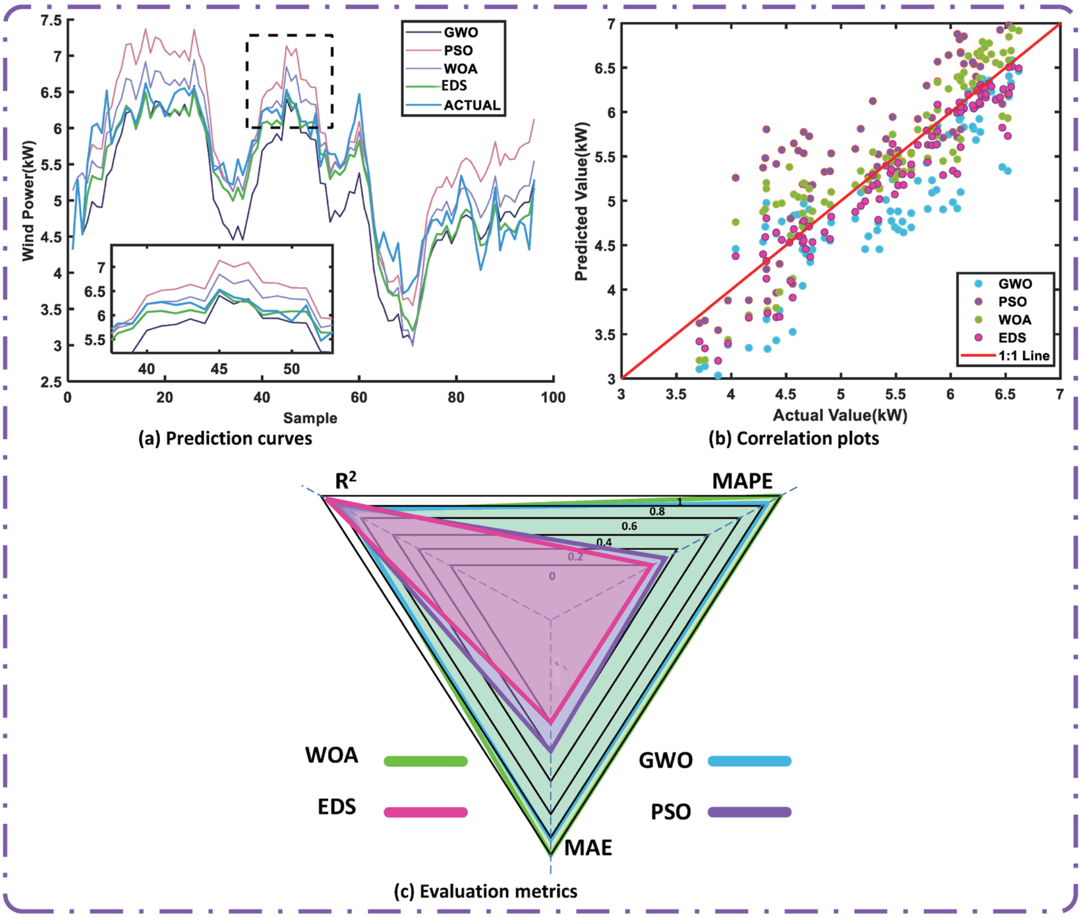

Fig. 7. (Color online) Wind power prediction for Case 3.

Table 8
Prediction indicators of different models for Case 3.

Models	Assessment criteria		
	RMSE	R^2 (%)	*MAPE* (%)
PSO-based model	0.54	78.70	8.34
WOA-based model	0.58	77.01	8.96
GWO-based model	0.37	84.37	5.91
EDS-based model	0.31	91.33	4.55

and the model's ability to extract key information for wind power prediction under adverse weather conditions.

To verify the effectiveness of the proposed model in midterm forecasting, data from the wind farm between February and April 2018 were utilized to compare its performance with that of the model proposed in the literature, LSTM-Attention, and Informer. With a minimum interval of one day, the datasets from February and March were employed as the training set to predict the wind power in April. The prediction results are presented in Table 9.

Table 9
Prediction indicators of different models for midterm forecast.

Models	Assessment criteria		
	RMSE	R^2 (%)	*MAPE* (%)
LSTM-Attention	0.36	91.59	9.50
Informer	0.37	90.79	9.68
LSTM-DFAT	0.30	94.13	7.63

Table 9 demonstrates that the proposed model also exhibits high accuracy in midterm forecasting. Compared with the Informer model, its *RMSE* and *MAPE* metrics decreased by 0.07 and 2.05, respectively. In contrast to the model's performance in short-term forecasting, the LSTM-DFAT model does not show a significant decrease in midterm forecasting metrics, indicating that LSTM-DFAT has excellent performance in both short-term and midterm forecasting.

6. Conclusions

Accurate wind power prediction contributes to improving grid stability and increasing wind power penetration. However, the inherent randomness and intermittency of wind power pose challenges to prediction accuracy. Therefore, in this study, we proposed a wind power prediction model that combines deep learning networks with directed attention mechanisms. The model initially preprocesses data by employing PCA to handle high-dimensional meteorological data, specifically temperature, humidity, wind speed, and wind direction, collected by FT-WQX5 sensors. Subsequently, a DFAT is utilized to enhance the identification of key information, and the integrated LSTM networks capture temporal characteristics. The main findings of this study are as follows.

(1) DFAT effectively enhances LSTM's ability to extract the most important information from highly nonlinear and multivariate data, addressing the challenges posed by wind power's uncertainty and intermittency in prediction tasks. This improves both the accuracy and generalization capabilities of the forecasting model. In Case Study 1, under various normal weather conditions, the LSTM-DFAT model achieved *RMSE* values below 0.8, R^2 values exceeding 90%, and *MAPE* values under 7% for wind power probability forecasting. These findings validate the LSTM-DFAT model's stable predictive performance and robustness in multivariate operational environments.

(2) The proposed EDS-PKO algorithm exhibits superior performance compared with conventional intelligent optimization algorithms when solving both unimodal and multimodal benchmark functions. Notably, for all test functions except F1 and F6, the EDS-PKO algorithm successfully converges to the theoretical optimal value of 0, with dynamic hyperparameter adjustment capabilities for LSTM-DFAT models based on environmental changes, enhancing wind power forecasting adaptability in complex scenarios.

(3) In Case 2, compared with the PSO, WOA, and GWO algorithms, the LSTM-DFAT model optimized using the EDS-PKO algorithm showed a significant improvement in *RMSE* values by 42.59, 46.55, and 16.21%, R^2 values by 16.04, 18.59, and 6.96%, and *MAPE* values by 3.97, 4.41, and 1.36%, respectively, demonstrating its superiority.

The main contributions and innovative achievements of this study are summarized as follows: (1) a wind power composite prediction model based on directed attention mechanisms and deep learning was proposed, (2) a DFAT was constructed, and (3) the EDS-PKO algorithm that integrates dynamic parameter control and dual-stage search mode was proposed.

Although the proposed model demonstrates excellent performance in short-term wind power prediction, there are limitations in two aspects: first, while the model shows good predictive performance under adverse weather conditions, the frequency and intensity of extreme weather events may vary with climate; second, the generalization ability of the proposed model and algorithm needs further improvement. Future research will address these challenges.

Acknowledgments

This work was supported by the Science and Technology Program of Chengde City, People's Republic of China [Project No. 202305B029].

References

1 K. Umino, M. L. L. Li, B. X. Ji, M. K. Lim, and M. L. Tseng: Appl. Soft Comput. **150** (2024) 111087. https://doi.org/10.1016/j.asoc.2023.111087
2 Global Wind Report 2024 (2024). https://gwec.net/global-wind-report-2024
3 Y. Wang, R. M. Zou, F. Liu, L. J. Zhang, and Q. Y. Liu: Appl. Energy **304** (2021) 117766. https://doi.org/10.1016/j.apenergy.2021.117766
4 Z. Zhang, X. Gao, Q. Chen, and Y. Yuan: Phys. Fluids **36** (2024) 055124. https://doi.org/10.1063/5.0203159
5 L. L. Li, B. X. Ji, G. C. Liu, J. P. Yuan, S. W. Tseng, M. K. Lim, and M. L. Tseng: Energy **294** (2024) 130905. https://doi.org/10.1016/j.energy.2024.130905
6 C. Y. Liu, X. M. Zhang, S. W. Mei, Z. Zhen, M. S. Jia, Z. Li, and H. Y. Tang: Appl. Energy **313** (2022) 118769. https://doi.org/10.1016/j.apenergy.2022.118769
7 P. Salgado, G. Igrejas, and P. Afonso: 13th APCA Int. Conf. Control and Soft Computing (CONTROLO), (Univ Azores, Ponta Delgada, PORTUGAL, 2018). https://doi.org/10.1109/CONTROLO.2018.8514268
8 B. Q. Lin and C. C. Zhang: Renewable Energy **179** (2021) 1565. https://doi.org/10.1016/j.renene.2021.07.126
9 K. Wang, X. Y. Tang, and S. M. Zhao: Expert Syst. Appl. **238** (2024) 121886. https://doi.org/10.1016/j.eswa.2023.121886
10 Aasim, S. N. Singh, and A. Mohapatra: Renewable Energy **136** (2019) 758. https://doi.org/10.1016/j.renene.2019.01.031
11 R. G. Kavasseri and K. Seetharaman: Renewable Energy **34** (2009) 1388. https://doi.org/10.1016/j.renene.2008.09.006
12 L. Yu, Y. X. Ma, Y. M. Ma, and G. X. Zhang: Sustainable Energy Technol. Assess. **49** (2022) 101794. https://doi.org/10.1016/j.seta.2021.101794
13 H. S. Dhiman, D. Deb, and J. M. Guerrero: Renewable Sustainable Energy Rev. **108** (2019) 369. https://doi.org/10.1016/j.rser.2019.04.002
14 Y. L. Liu and J. Y. Wang: Appl. Energy **312** (2022) 118729. https://doi.org/10.1016/j.apenergy.2022.118729
15 J. H. Pang and S. Dong: Energy Convers. Manage. **307** (2024) 118343. https://doi.org/10.1016/j.enconman.2024.118343
16 T. Y. Ji, J. Wang, M. S. Li, and Q. H. Wu: Energy Convers. Manage. **254** (2022) 115196. https://doi.org/10.1016/j.enconman.2021.115196
17 V. K. Rayi, S. P. Mishra, J. Naik, and P. K. Dash: Energy **244** (2022) 122585. https://doi.org/10.1016/j.energy.2021.122585
18 Y. Y. Shao, J. Z. Wang, H. P. Zhang, and W. G. Zhao: Appl. Math. Modell. **100** (2021) 780. https://doi.org/10.1016/j.apm.2021.07.024
19 W. L. Fu, K. Wang, J. W. Tan, and K. Zhang: Energy Convers. Manage. **205** (2020) 112461. https://doi.org/10.1016/j.enconman.2019.112461

20 B. Yang, L. N. Zhong, J. B. Wang, H. C. Shu, X. S. Zhang, T. Yu, and L. M. Sun: J. Cleaner Prod. **283** (2021) 124628. https://doi.org/10.1016/j.jclepro.2020.124628
21 M. Ding, H. Zhou, H. Xie, M. Wu, Y. Nakanishi, and R. Yokoyama: Neurocomputing **365** (2019) 54. https://doi.org/10.1016/j.neucom.2019.07.058
22 Z. F. Liu, X. R. Chen, Y. H. Huang, X. F. Luo, S. R. Zhang, G. D. You, X. Y. Qiang, and Q. Kang: Energy **303** (2024) 131947. https://doi.org/10.1016/j.energy.2024.131947
23 S. Wang, B. Li, G. Z. Li, B. Yao, and J. Z. Wu: Appl. Energy **292** (2021) 116851. https://doi.org/10.1016/j.apenergy.2021.116851
24 Y. X. Guo, Y. Li, X. B. Qiao, Z. Y. Zhang, W. F. Zhou, Y. J. Mei, J. J. Lin, Y. C. Zhou, and Y. Nakanishi: IEEE Trans. Smart Grid **13** (2022) 3481. https://doi.org/10.1109/TSG.2022.3173964
25 B. R. Xiong, L. Lou, X. Y. Meng, X. Wang, H. Ma, and Z. X. Wang: Electr. Power Syst. Res. **206** (2022) 107776. https://doi.org/10.1016/j.epsr.2022.107776

About the Authors

Zhixin Yu graduated from Hebei University of Engineering, China, in 2022. He is an engineer at Hebei Shengde Infrastructure Construction Development Co. (yuzhixin1989@126.com)

Huaichang Liu received his B.S. degree from the City College of Hebei University of Technology, China, in 2007. He is an engineer at Hebei Shengde Infrastructure Construction Development Co. (lhch831118@126.com)

Kunyu Liu graduated from Hebei University of Science and Technology, China, in 2023. He is an engineer at Hebei Shengde Infrastructure Construction Development Co. (lku8908@126.com)

Jingguo Ma received his B.S. degree from Hebei University of Technology, China, in 1991. He is an engineer at Hebei Shengde Infrastructure Construction Development Co. (mjg19980101@126.com)

Bingxiang Ji is a Ph.D. student of electrical engineering in the School of Electricity, Hebei University of Technology. His research interests include power systems, new energy sources, and integrated energy systems. (202311401014@hebut.edu.cn)

Ling-ling Li received her M.S. degree in control theory and control engineering in 2001 and her Ph.D. degree in electrical machinery and appliances in 2004 from Hebei University of Technology. She is currently a professor at the School of Electrical Engineering, Hebei University of Technology, and a permanent member of the State Key Laboratory of Reliability and Intelligence of Electrical Equipment. Her research interests include power systems, new energy, and electrical reliability.
(lilingling@hebut.edu.cn)