IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 4, 2021, accepted June 26, 2021, date of publication June 30, 2021, date of current version July 8, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3093646

Short-Term Wind Power Prediction Based On Particle Swarm Optimization-Extreme Learning Machine Model Combined With Adaboost Algorithm

GUOQING AN[1,2], ZIYAO JIANG[1], XIN CAO[3], YUFEI LIANG[1], YUYANG ZHAO[1,2], ZHENG LI[1,2], (Member, IEEE), WEICHAO DONG[1,2], AND HEXU SUN[1,2], (Senior Member, IEEE)

[1]School of Electrical Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China
[2]Hebei Engineering Laboratory of Wind Power/Photovoltaic Coupling Hydrogen Production and Comprehensive Utilization, Shijiazhuang 050018, China
[3]Hebei Construction & Investment Group New Energy Company Ltd., Shijiazhuang 050051, China

Corresponding author: Hexu Sun (sunhxhb@outlook.com)

This work was supported in part by the Project of Key Research and Development Plan of Hebei Province under Grant 20314501D and Grant 19214501D.

ABSTRACT As the proportion of wind power in the world's electricity generation increases, improving wind power prediction accuracy is vital for making full use of wind energy and ensuring the safe and stable operation of the power grid. Given the uncertainty and volatility of wind power and the weak generalization ability of the current wind power prediction models, we propose a wind power prediction model that combines Adaboost algorithm with extreme learning machine optimized by particle swarm optimization (PSO-ELM). First, particle swarm optimization is used to optimize the initial thresholds and input weights of the ELM to obtain the PSO-ELM basic prediction model. Then, combined with the Adaboost algorithm, a series of PSO-ELM weak predictors with input weights and thresholds optimized by PSO and containing different hidden layer nodes are composed. Finally, each weak predictor is weighted and fused into a strong prediction model of wind power, and the final prediction results are output. In this paper, the Adaboost-PSO-ELM model is verified by a wind turbine's measured data in Turkey. The prediction indicators are compared with the current wind power prediction methods including optimized neural networks and ensemble learning models. The results show that the Adaboost-PSO-ELM wind power prediction model has higher accuracy and better generalization ability.

INDEX TERMS Adaboost algorithm, extreme learning machine, optimization algorithm, wind power prediction.

NOMENCLATURE

BP	Back Propagation Neural Network
ELM	Extreme Learning Machine
GA	Genetic Algorithm
GPR	Gaussian Process Regression
GWEC	Global Wind Energy Council
KELM	Kernel Extreme Learning Machine
LSTM	Long Short-Term Memory
NWP	Numerical Weather Report
PSO	Particle Swarm Optimization
SLFN	Single-Hidden Layer Feedforward Neural Network
SVM	Support Vector Machine

The associate editor coordinating the review of this manuscript and approving it for publication was Seyyed Ali Pourmousavi Kani.

I. INTRODUCTION

At present, wind power has become the most mature power generation technology in all renewable energy business models [1]. The latest report of the Global Wind Energy Council (GWEC) shows that the total wind power capacity in Asia Pacific is now nearly 347 GW, making it the region with the most wind power capacity worldwide [2]. However, wind power generation's randomness poses a challenge to the power grid's safe dispatch and stable operation. Therefore, accurate wind power prediction is of great significance for reducing grid dispatching costs and improving system performance. According to time scale, wind power prediction can be classified into long-term, medium, short-term and ultra-short-term prediction. Long-term prediction with annual timeliness are used in wind farm planning and annual power generation plans; mid-term prediction are time-limited by weeks or months to determine maintenance plans. The accuracy requirements of these two prediction methods are not strict, and there are few studies on this aspect. On the other hand, the short-term prediction refer to the prediction of wind power within 3 days. And the ultra-short-term prediction means to predict the wind power in the next 10 minutes to 4 hours [3]. They can help in the optimization of frequency modulation and spinning reserve capacity, as well as the economic load dispatch, which makes it a hot spot in the wind power industry. Recent studies have mostly focused on the prediction of the wind power time series. In [4], Zhou *et al.* proposed to use K-means clustering to process historical wind power data to acquire multiple training sets, and then use long short-term memory (LSTM) neural network to predict the future wind power. Yan and Wu [5] used wavelet transform to deconstruct the wind power time series to obtain multiple sub-sequences, input the kernel extreme learning machine (KELM) for prediction, and then integrate the output results to obtain the final prediction value. Lee *et al.* [6] proposed to integrate three ensemble learning models, including boosted trees, random forest and generalized random forest, to realize the prediction of wind power. Sun and Zhao [7] used VMD method to process historical power series, and then used LSTM to predict and fuse sub-sequences. Shi *et al.* [8] proposed to use the wavelet decomposition method to process the time series of wind power and then uses SVM to predict wind power in the last day of a week. These time series-based methods are effective in ultra-short-term prediction, but for short-term time scale prediction, their robustness is insufficient and it is difficult to achieve satisfactory results.

Methods based on physical models and statistical models are the mainstream methods for short-term wind power prediction [9]. The core of the physical-model-based method is to establish the dynamic and thermodynamic equations corresponding to the evolution of the meteorology, and then combine the actual latitude, altitude and terrain of the wind farm to predict the wind speed to finally realize the wind power prediction through the wind speed-power fitting curve [10]. This prediction model is more complicated, with a large amount of calculation, and is easily affected by incorrect initial information [11]. The statistical-model-based method uses training samples to build a model, which can be combined with the training model to obtain the predicted wind power according to the new input measurement values [12]. Because it establishes a nonlinear relationship between the input and the output, the model can be applied to problems with unclear internal mechanisms and have broader applicability in the field of wind power prediction [13]. In recent years, scholars worldwide have carried out much research on the realization of wind power forecasting by constructing statistical models. Among them, the two methods of support vector machine (SVM) and neural network have achieved satisfactory prediction results [14]. In terms of realizing wind power prediction based on the SVM method, Ning and Liu [15] used particle swarm algorithm to optimize the penalty factor and kernel parameters of SVM and then used historical data as training samples to train the optimized SVM model. In terms of realizing wind power prediction based on neural network method, Li and Mao [16] put forward a method using two-day historical climate data and wind power data to train BP neural network, and then predict the ultra-short-term wind power in the next 4 hours based on the numerical weather forecast. Wang *et al.* [17] proposed to use empirical mode decomposition to reconstruct the time series of wind speed and input the obtained signal components into the BP neural network optimized by genetic algorithm to realize the prediction of future wind speed. And the wind power prediction is achieved by fitting the wind speed-power curve. But there are some problems with these popular methods: First, the training time of SVM is too long, and there's much difficulty to implement large-scale training samples; second, the BP neural network requires many parameters and thereby cause sample dependence, which makes it fall into local optimality frequently.

The Extreme Learning Machine (ELM) is an improvement to the traditional single-hidden layer feedforward neural network (SLFN), and it has made much progress in recent years. ELM randomly generates the hidden layer neurons' threshold and the connection weights between the input layers, which does not need to be adjusted during training. Its structure is not complicated and it has the advantages of fast training speed, and better generalization ability [18] and has recently been used in the research of wind power prediction. However, ELM input parameters' randomness will inevitably affect the prediction results. To improve the prediction accuracy of ELM, some scholars have applied different optimization algorithms to optimize the input parameters of ELM. Wang *et al.* [19] used genetic algorithm to optimize the input weights of ELM; Zhai and Ma [20] and Tan *et al.* [14] used artificial fish swarm algorithm and salp swarm algorithm to optimize the initial input weights and thresholds of the ELM, respectively. Although these optimization algorithms have improved the prediction accuracy of the model to a certain extent, in some cases, there will be over-fitting phenomena, which will make the model fall into the local optimum and affect the generalization ability of the prediction model [14]. The Adaboost algorithm is an integrated learning algorithm proposed by Schapire and Freund [21]. It aims to aquire and adjust all training samples' weight through repeated iterative training of the model. Each iteration forms a weak predictor and then combines these weak predictors into a strong predictor through weighting, thereby improving the overall model's performance. The Adaboost algorithm can be combined with a variety of learning algorithms. In [22], Xiang *et al.* combined Adaboost with BP neural network to predict the future tax; In [23], the particle swarm optimization SVM was combined with Adaboost to forecast the water inrush from the coal seam floor. Xiang and Zhu [22], Wen and Yu [23] verified that the Adaboost algorithm improves the generalization ability of the model and prevents the training process of the learning algorithm from falling into the local optimum.

Combining the Adaboost algorithm with the extreme learning machine, the Adaboost-PSO-ELM wind power prediction model is proposed. Particle swarm optimization is applied to find the best suited input weights and thresholds of the model. And then we use the PSO-ELM model as a weak predictor. The number of each weak predictor's hidden layer nodes is set in a validated interval. Then we use the Adaboost algorithm to integrate them into a strong predictor of wind power to further improve the model's prediction's accuracy and ability of generalization. Finally, the actual data is used to train the proposed model. The prediction results are compared with the results of the existing models in order to verify the superior performance of the proposed wind power prediction model.

The remainder of this article is arranged as follows: In Section II, we introduce the principles of ELM, particle swarm optimization algorithm, and Adaboost integrated learning algorithm in detail, and the Adaboost-PSO-ELM wind power prediction model is proposed; In Section III, the Adaboost-PSO-ELM wind power prediction model is trained and output prediction results based on smaller samples and larger samples, and the results are compared with the results of several current mainstream models. Through multi-index error analysis, the validity and applicability of the method proposed in this article are verified; In Section IV, we draw some conclusions.

II. BASIC THEORIES

A. WIND POWER PREDICTION METHOD BASED ON EXTREME LEARNING MACHINE

Aiming at the traditional single-hidden layer feedforward neural network (SLFN) easy to fall into the local extremum, with slow training speed and other problems [24], Huang *et al.* proposed an extreme learning machine (ELM) method. This method can randomly generate the threshold value of the hidden layer neurons and the connection weight between the input layer and the hidden layer. There is no need for iterative adjustment during the training process, so it possesses a faster training speed and better nonlinear fitting ability.

In the following, the wind power prediction method based on the extreme learning machine will be explained by taking wind power prediction through wind speed and wind direction as an example. Given M training samples of wind power $\left(x\_train_j, t\_train_j\right)$, $j = 1, 2, \ldots, M$. Among them, x_train_j represents the input information of the training sample, including the measured wind speed and wind direction

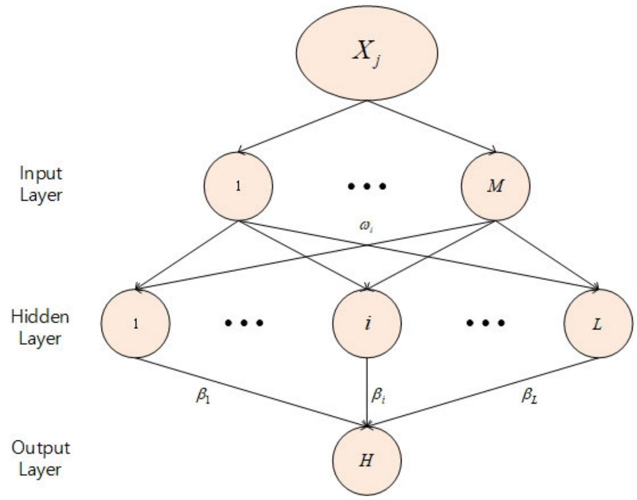

FIGURE 1. Structure diagram of ELM.

data, t_train_j is the output data of the training sample, corresponding to the wind power data. Suppose the network contains L hidden layer neurons, activation function ReLU:

$$g(x) = \begin{cases} x, x > 0 \\ 0, x \leq 0 \end{cases}. \tag{1}$$

Then the output of the network model of the j-th sample is expressed as

$$\sum_{i=1}^{L} \beta_i g(\omega_i x\_train_j + b_i), \quad i = 1, 2, \ldots, L \tag{2}$$

where ω_i is the connection weight between the input node and the i-th hidden layer node; β_i is the weight between the output node and the i-th hidden layer node; b_i is the threshold of the i-th hidden layer node.

The hidden layer output matrix of the ELM can be expressed as:

$$H = \begin{pmatrix} g(\omega_1 x\_train_1 + b_1) & \ldots & g(\omega_L x\_train_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(\omega_1 x\_train_M + b_1) & \cdots & g(\omega_L x\_train_M + b_L) \end{pmatrix} \tag{3}$$

where $\beta = [\beta_1, \beta_2, \ldots, \beta_L]^T$, $T = [t\_train_1, \ldots, t\_train_M]^T$ The core of the wind power prediction method based on the extreme learning machine is to solve the output weight vector $\hat{\beta}$, which makes

$$\left\| H\hat{\beta} - T \right\| = \min_{\beta} \|H\beta - T\| \tag{4}$$

whose least square solution is:

$$\hat{\beta} = H^{\dagger} T \tag{5}$$

in which $H^{\dagger}$ is the Moore-Penrose generalized inverse matrix of H.

In this way, if the short-term prediction data of wind speed and wind direction are known, and the $\hat{\beta}$ and $H^{\dagger}$ obtained through the training sample are combined, the short-term prediction value of wind power can be obtained.

B. PARAMETER OPTIMIZATION OF EXTREME LEARNING MACHINE BASED ON PARTICLE SWARM

The particle swarm algorithm's main idea is to initialize each potential optimal solution of the optimization problem to a particle with an initial velocity, position, and fitness value. Each particle updates its speed and position based on individual flight experience and group experience in each iteration of the optimization process. Moreover, the fitness value is used to judge the current position. The individual and global optimal solutions are approximated to obtain the final individual best solution *Pbest* and global best solution *Gbest* [25]. The process can be described as

$$V_i^{k+1} = \lambda V_i^k + c_1 r_1^k \left(Pbest_i^k - X_i^k \right) + c_2 r_2^k (Gbest_i^k - X_i^k) \tag{6}$$

$$X_i^{k+1} = X_i^k + V_i^{k+1} \tag{7}$$

in which $k = 1, 2, \ldots, K$, and K is the maximum iterations for the PSO process; λ is the inertia factor that regulates the particle's speed; c_1 and c_2 are the acceleration factors that regulate particle speed. In the k-th iteration, r_1^k and r_2^k are randomly distributed in [0,1]; X_i^k is the position of the particle; V_i^k is the velocity of the particle; $Pbest_i^k$ and $Gbest_i^k$ are the individual optimal position and the global optimal position of the particle.

Considering that the weight ω_i and the threshold b_i will directly affect the wind power prediction results based on the ELM model, the PSO algorithm can be integrated with the ELM method to achieve the global optimization of ω_i and b_i. In the following, we describe the details of the process:

Step 1: Initialize the maximum number of iterations K, inertia factor λ, acceleration factor c_1, c_2 and target error $err_{\min}$.

Step 2: Initialize the population. Determine Z as the particles' number. Assign a set of information to every particle about the weight ω_i and threshold b_i of the ELM model. The mean square error (MSE) of training samples' predicted value in the ELM model is taken as the fitness value ψ of the optimization process. All particles' fitness value in the population is calculated, and the initial values of *Pbest* and *Gbest* is determined by the information of the particle with the smallest ψ

Step 3: Update the information of the population. According to formula (5) and formula (6), all particle positions and velocities are updated. If the current position information of a specific particle after the update has a smaller ψ value than *Pbest* before the update, replace the position information with *Pbest*, otherwise *Pbest* remains unchanged; If the best position of the population after the update has a smaller ψ value than the one before the update, then the position information is replaced with *Gbest*, otherwise *Gbest* remains unchanged.

Step 4: When $\psi < err_{\min}$ or the number of iterations reaches K, the iteration stops. Meanwhile, the optimized combination of ω_i and b_i in the ELM model equals the particle information in correspondence with *Gbest*.

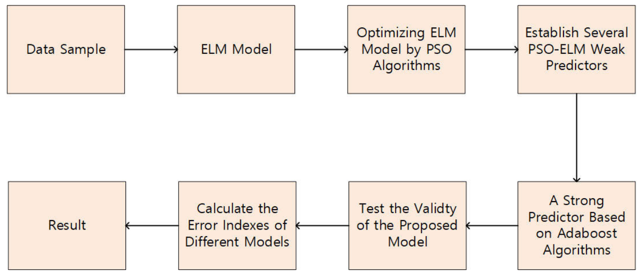

FIGURE 2. The flow chart of the study.

C. ADABOOST-PSO-ELM WIND POWER PREDICTION MODEL

The Adaboost algorithm is implemented by changing the weight of the data. First, given the error threshold φ, the training samples are used in the weak predictors' training. For samples with an error greater than φ in the prediction result, we increase its weight, otherwise reduce its weight. Moreover, the weight of each weak predictor is calculated according to the sample weight, and then a strong predictor is formed [17].

In addition to weights and thresholds, the ELM model's predictive performance is also affected by the hidden layer nodes' number L, and this effect is uncertain and nonlinear for different sets of data. The hidden layer nodes' number is related to the input and output dimensions. When M represents the training samples' number, its number must be less than $M - 1$. Suppose this condition was not met. In that case, there is no guarantee that each node corresponds to a sample after averaging. And this phenomena may weaken the model's generalization ability. If there are too many nodes, the prediction model may be overfitted. In the existing literature, L is often set to a fixed value through experience, which has many limitations.

Besides, in the PSO optimization process, the ELM model may still fall into the local optimum. To boost the model's generalization ability and avoid the PSO-ELM wind power prediction model falling into the local optimum and overfitting's risk, this paper proposes the Adaboost-PSO-ELM wind power forecasting method. This method regards the PSO-ELM model as a weak predictor. Through the evaluation of the prediction results of the validation set, the number of hidden layer nodes L of each weak predictor is set in the interval with the best performance. Based on the Adaboost method, multiple weak predictors with different weights are obtained through repeated iterative training with a set of training samples, and finally weighted and fused into a strong predictor to complete wind power prediction, to further reduce the prediction error of the ELM model and improve the model's generalization performance. The details are as follows:

Step 1: Initialization. Initialize $\varepsilon_0 = 0$, and set the error threshold φ. Suppose the number of the model's training samples is M, set the number of weak predictors to Z, and the weight coefficient of the training samples of the first weak predictor is $D_z(i)$, $z = 1, 2, \ldots, Z$. Initialize the distribution weight of the first weak predictor training sample as

$$D_1(i) = 1/M, i = 1, 2, \cdots, M \tag{8}$$

Step 2: The creation of a weak predictor. Input the M training samples with the weight coefficient distribution $D_1(i)$ into the first PSO-ELM wind power forecasting model, and take it as a weak predictor. The number of hidden layer nodes of the weak predictor is distributed in the validated interval. $D_z(i)$ are the weight coefficient of the samples. They are not used as input data to participate in the training of the PSO-ELM model. Its significance is to find the weight coefficient a_z of each weak predictor.

Step 3: Determine the weight coefficient a_z of the weak predictor. Input the training samples (wind speed and direction) to the weak predictor to obtain the output (wind power) of the training samples. We compare the prediction output of the training sample with the actual power, and the prediction error sequence of the z-th weak predictor is calculated as:

$$e_z(i) = \|f_z(x\_train_i) - t\_train_i\| \tag{9}$$

where $i = 1, 2, \ldots, M, z = 1, 2, \ldots, Z$. And $f_z(x\_train_i)$ represents the predicted wind power value of the z-th weak predictor when the x_train_i of the i-th training sample is used as input and t_train_i represents the actual wind power value of the i-th training sample.

According to the value of $e_z(i)$, the weight coefficients of the samples whose prediction error is greater than the threshold φ are accumulated, and then the weight coefficient a_z of the z-th weak predictor is obtained:

$$\varepsilon_z = \sum_{i:e_i>\phi} D_z(i) \tag{10}$$

$$a_z = \frac{1}{2}\ln\left(\frac{1-\varepsilon_z}{\varepsilon_z}\right) \tag{11}$$

Step 4: Update sample weights. According to the sample weight coefficient $D_z(i)$ of the z-th weak predictor, we use the following formula to determine the sample weight coefficient $D_{z+1}(i)$ of the next weak predictor:

$$D_{z+1}(i) = \frac{D_z(i)}{B_z} \times \begin{cases} 1, e_1(i) > \varphi \\ \varepsilon_z^{\mu}, e_1(i) \leq \varphi \end{cases} ; z = 1, 2, \cdots, Z-1 \tag{12}$$

In the formula, μ is the power coefficient, the value is [26], and set to 2; the B_z in formula (11) is the normalization factor whose function is to keep the sum of the distribution weights as 1:

$$B_z = \sum_{i=1}^{M} D_z(i) \tag{13}$$

Step 5: Repeat steps 2 to 4 in a loop until the training of the Z-th weak predictor is completed.

Step 6: Build a strong predictor. According to the Z weak predictors and their weight coefficients, the strong predictor function $F(x)$ is obtained by the following formula:

$$F(x) = \sum_{z=1}^{Z} a_z f_z(x) \tag{14}$$

III. CASE ANALYSIS

This paper uses the wind power public data set provided by Kaggle as the training and test data of the Adaboost-PSO-ELM wind power prediction model. This data set comes from a SCADA system in a wind turbine in Yalova, Turkey. Active power (kW), average wind speed (m/s) and absolute wind direction (°) for all twelve months of 2018 are contained in the data set recorded by the SCADA system. The measurement interval is 10 minutes, and a total of 50531 sets of valid data are available. The first five sets of data are as Table 1 shows:

TABLE 1. The first five sets of data in the data set.

Sample	Wind Speed(m/s)	Wind Direction(°)	Wind Power(kW)
1	5.31	260.0	380.05
2	5.67	268.6	453.77
3	5.22	272.6	306.38
4	5.66	271.3	419.65
5	5.58	265.7	380.65

A. MODEL PERFORMANCE EVALUATION

In order to compare the reliability and prediction performance of neural networks and ensemble learning models involved in this article, mean square error (MSE), root mean square error (RMSE), mean absolute percentage error (MAPE), mean absolute error (MAE), mean bias error (MBE), relative mean bias error (RMBE) and coefficient of determination (R^2) are used as evaluation indicators for model errors. The calculation formulas are:

$$E_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(t\_predict_i - t\_test_i)^2 \tag{15}$$

$$E_{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t\_predict_i - t\_test_i)^2} \tag{16}$$

$$E_{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{t\_predict_i - t\_test_i}{t\_test_i}\right| \tag{17}$$

$$E_{MAE} = \frac{1}{N}\sum_{i=1}^{N}|t\_predict_i - t\_test_i| \tag{18}$$

$$E_{MBE} = \frac{1}{N}\sum_{i=1}^{N}(t\_predict_i - t\_test_i) \tag{19}$$

$$E_{RMBE} = \frac{\sum_{i=1}^{N}(t\_predict_i - t\_test_i)}{\sum_{i=1}^{N} t\_predict_i} \cdot 100 \tag{20}$$

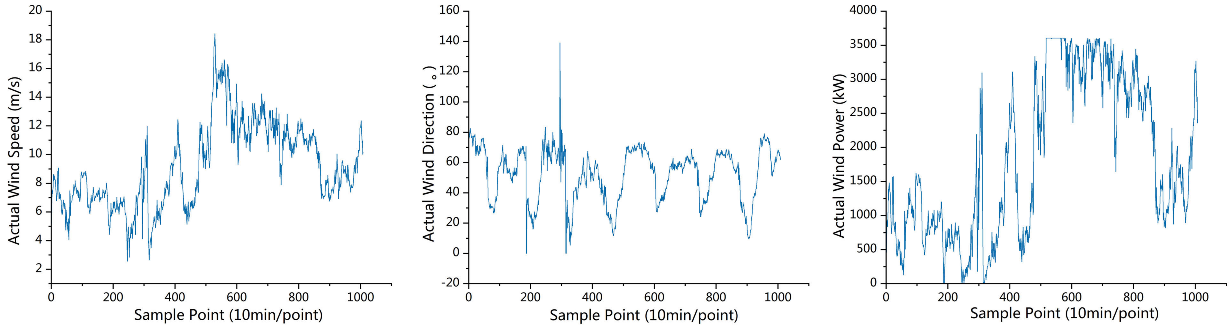

FIGURE 3. **Measured wind speed, wind direction and wind power (7 consecutive days).**

$$R^2 = \frac{\sum_{i=1}^{N} (t\_predict_i - \overline{t\_test})^2}{\sum_{i=1}^{N} (t\_test_i - \overline{t\_test})^2} \tag{21}$$

in which N represents the number of predicted samples, $t\_predict_i$ corresponds to the value of predicted wind power, and t_test_i to the value of actual wind power. The predictive performance of the model is positively correlated with R^2, and negatively correlated with other indicators listed above.

B. SMALL SAMPLE WIND POWER PREDICTION MODEL BASED ON EXTREME LEARNING MACHINE

To evaluate the model's predictive performance, this paper uses wind speed and wind direction as input data and wind power as output data. First, samples with negative power values in the data set are deleted. We select 1008 sets of valid data in 7 days from August 15 to August 21 as the data set of this model. The data set is divided into three parts: training set, validation set and test set, and the model construction, optimal parameter selection and performance evaluation are realized respectively, and the ratio is 6:2:2. The time series of wind speed, wind direction, and wind power in 7 days are shown in Fig. 3. In order to improve the prediction accuracy and speed up the optimization of PSO, we use Min-Max normalization to preprocess the data, the normalization function is

$$x_{im} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \tag{22}$$

and the data is mapped to [0,1] before prediction.

The extreme learning machine (ELM) is used to process the data set. Through the performance of the validation set, which is shown in Fig.4, we evaluate the prediction performance of the ELM model under different activation functions with different number of hidden layer nodes. And the ReLU function is selected as the activation function of ELM, and the number of nodes is set to 30. The prediction results and actual values are as Fig. 5 shows.

It is illustrated in the Fig.5 that the changing trend of the ELM model's prediction results is basically in line with the actual values, but the prediction error is relatively large.

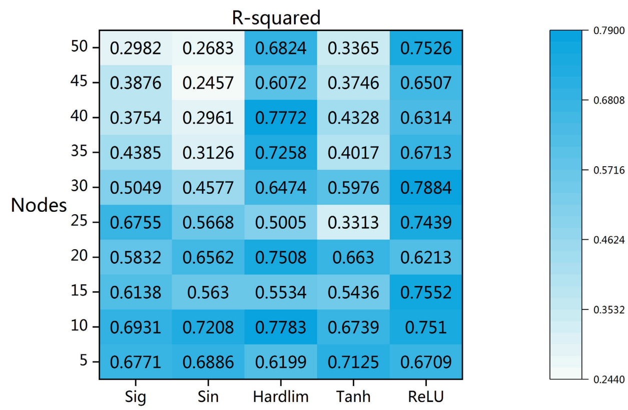

FIGURE 4. **The prediction performance of ELM model with different activation functions and hidden layer nodes.**

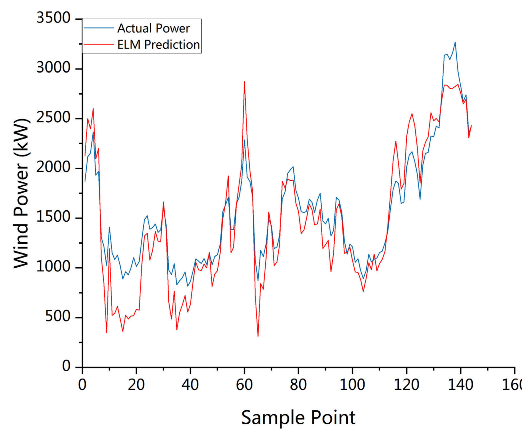

FIGURE 5. **Comparison of ELM prediction results and actual power.**

The reason is that the input weights and thresholds of the traditional ELM model are randomly generated by the program. Therefore, there are still some key points for research in the optimization of input weights and thresholds.

To solve the problems above, particle swarm optimization is used to optimize the input weights and thresholds of the extreme learning machine. The initial parameters in this paper are set as follows: the number of input layer nodes is set to 2. According to the validation results that Fig. 6 shows, the number of hidden layer nodes is set to 95, the number

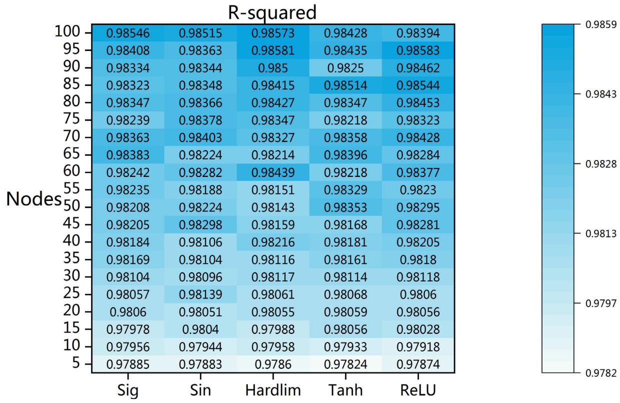

FIGURE 6. The prediction performance of PSO-ELM model with different activation functions and hidden layer nodes.

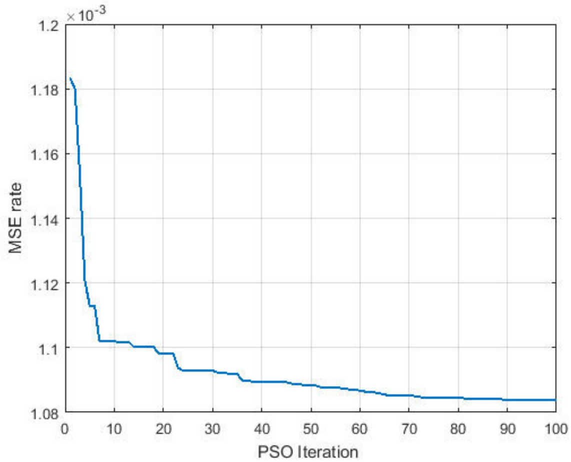

FIGURE 7. Fitness changes during PSO iteration.

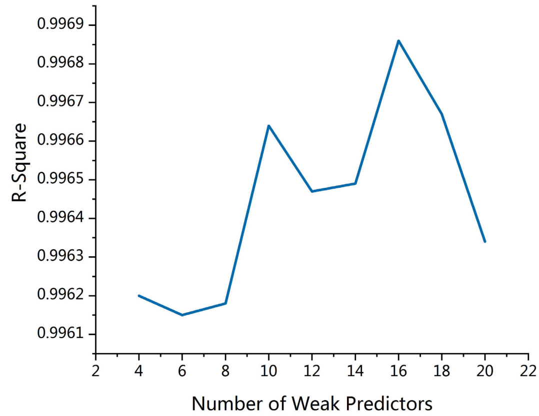

FIGURE 8. The coefficient of determination of Adaboost-PSO-ELM model with different number of weak predictors.

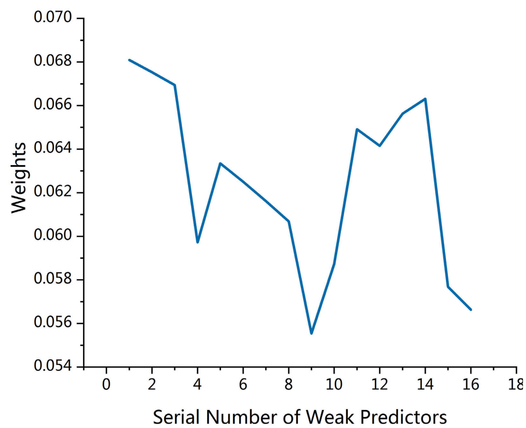

FIGURE 9. Weights of the weak predictors.

of output layer nodes is set to 1, ReLU is selected to be the activation function, and the maximum iterations of the particle swarm is 100, learning factor $c_1 = 1.8$, $c_2 = 1.2$, inertia factor $w = 1$, population size is set to 100. The ψ (MSE) of training samples is selected to be the optimization's fitness function. In Fig. 6 shows the change of particle fitness in the iterative process of PSO.

It can be seen from Fig. 7 that in the first ten iterations, the fitness of the particles has changed significantly. In the subsequent iterations, the changes in the fitness of the particles tend to be slight, and it starts to converge around the 90th iteration. Substitute the ELM weights and thresholds obtained from the PSO optimization into the ELM wind power prediction model. The comparison of the ELM and PSO-ELM methods' predictive indicators is shown in Table 2, and the comparison of the output results with the actual values of wind power is shown in Fig. 10.

Table 2 shows that all indexes in the PSO-ELM wind power prediction model involved in the table have been significantly improved compared to the ELM model. Among them, the MAE value, which directly represents the average deviation between the predicted data and the actual data, has been reduced by 79.55%, and the R^2, which represents the fitting performance of the regression model, has increased by 32.50%, which indicates that the prediction performance of the PSO-ELM model is greatly improved compared with ELM model, and the prediction results are more reliable. But there are still some deviations in a small number of the samples (20-25 sets and 90-100 sets in Fig. 10). This is mainly due to that the PSO optimization process cannot completely avoid the risk of the population falling into a local optimum. In addition, the number of hidden layer nodes determined by cross-validation may not be optimal when predicting different samples, which will affect the generalization performance of the prediction model. There are still some key points for improvement in the method of PSO optimized ELM.

To overcome these difficulties, this paper proposes the Adaboost-PSO-ELM wind power prediction model. As an improvement to the traditional PSO-ELM model, the Adaboost algorithm is applied to fuse multiple PSO-ELM weak predictors. We set the hidden layer nodes' number in a the interval with the best prediction performance for each weak predictor and conduct multiple times of training to further improve the generalization and prediction performance of the PSO-ELM model and avoid the PSO optimization from falling into the local optimum. After the validation

TABLE 2. Comparison of ELM and PSO-ELM predictive indicators.

	MSE	RMSE	MAPE	MAE	MBE	RMBE	R^2
ELM	64856	254.6683	0.1447	214.2717	-84.2765	-5.6772	0.7884
PSO-ELM	4549	67.4460	0.0372	50.5710	0.4621	0.2950	0.9857

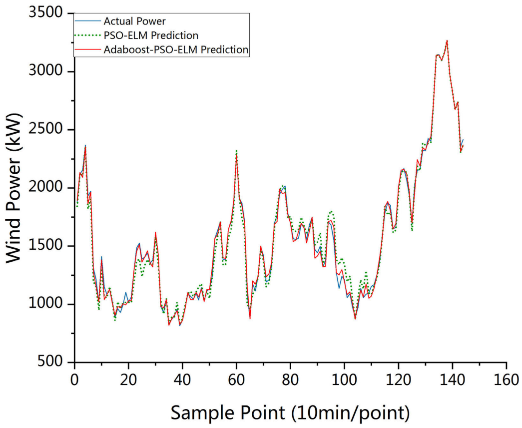

FIGURE 10. Comparison of PSO-ELM,Adaboost-PSO-ELM prediction results and actual power.

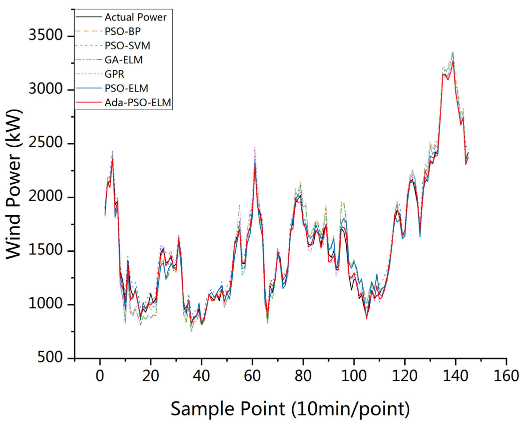

FIGURE 11. Comparison of prediction results with optimized neural network prediction methods.

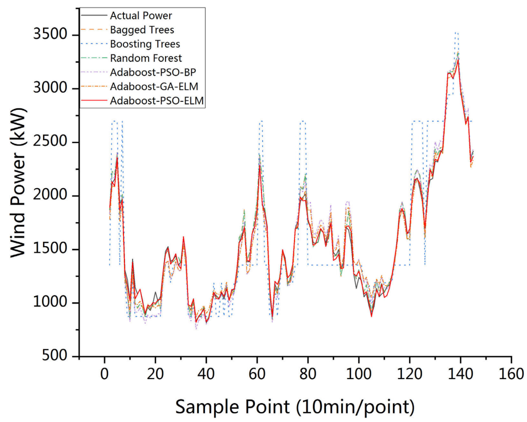

FIGURE 12. Comparison of prediction results with ensemble learning prediction methods.

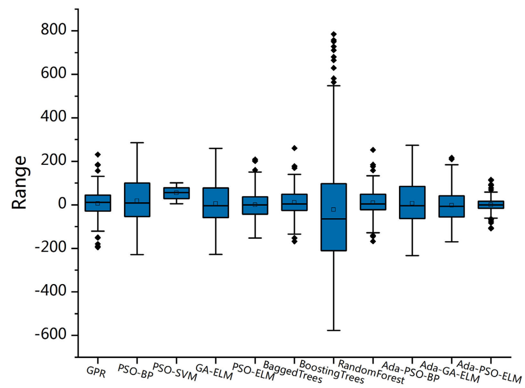

FIGURE 13. Comparison of prediction error box diagram with current prediction methods.

shown in Fig. 8, the number of weak predictors $Z = 16$, and the error threshold is set to the median of the absolute value of the PSO-ELM prediction error $\varphi = 40.65$. To ensure the wide applicability of the model, the number of hidden layer nodes is distributed randomly between 70-95, where the value of R^2 is the highest. The final weight distribution of the predictors is shown in Fig. 9, and the prediction result after the fusion of 16 weak predictors is also shown in Fig. 10.

From Fig. 10, it can be seen that in some samples where the prediction performance of the PSO optimized model is not satisfying (20-25 sets and 90-100 sets), the Adaboost algorithm can modify the prediction results of these samples by reducing the weights of them. It makes the prediction ability of the model proposed in this paper more accurate, and its overall prediction performance is also better than that of the PSO-ELM model. Besides, from the weight of each weak predictor in Fig. 9, it can be seen that the prediction performance of the model is not only affected by the number of hidden layer nodes, but also by the results of PSO optimization. The weight distribution of Adaboost effectively avoids the risk of particles falling into the local optimum in a single PSO optimization, thereby improves the overall prediction performance of the model.

In order to further verify the performance of the Adaboost-PSO-ELM method, several optimized neural networks and ensemble learning models including PSO-BP [27], PSO-SVM [28], GA-ELM [29], Gaussian Process Regression, Random Forest and Bagged Trees [6] are selected for comparative analysis. The index comparison of prediction error is shown in Table 3. The prediction results and the error box diagram are illustrated in Fig. 11, Fig. 12, and Fig. 13.

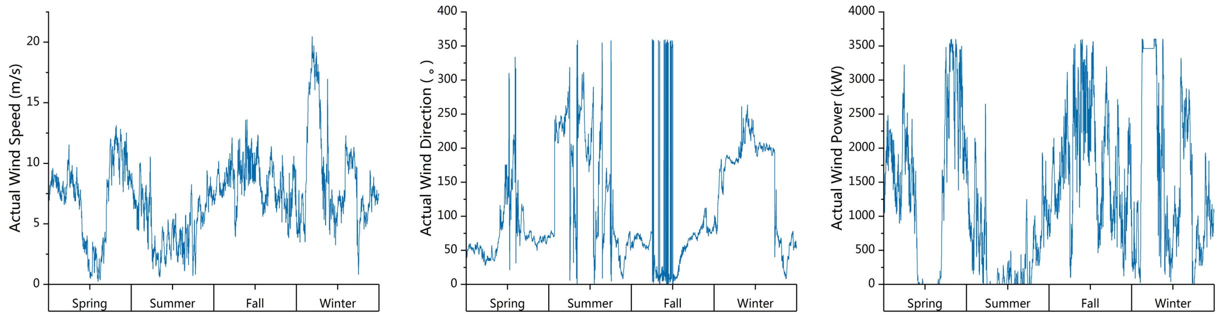

FIGURE 14. Measured wind speed, wind direction and wind power in different seasons (3.5 consecutive days each).

TABLE 3. Comparison of error indexes and training time with current prediction methods.

	MSE	RMSE	MAPE	MAE	MBE	RMBE	R^2	Traing Time/s
GPR	4657.3	68.2443	0.0345	50.3287	6.3316	0.4020	0.9853	56.912
PSO-BP	10946	104.6261	0.0598	85.0930	19.1548	1.2063	0.9655	19.655
PSO-SVM	3770.5	61.4045	0.0397	54.5305	54.5305	3.3593	0.9855	6357.317
GA-ELM	9197	95.9009	0.0559	76.8723	5.9521	0.3780	0.9710	49.833
PSO-ELM	4549	67.4460	0.0372	50.5710	0.4621	0.2950	0.9857	146.318
BaggedTrees	4408	66.3929	0.0330	48.7782	10.4678	0.6628	0.9861	45.136
BoostingTrees	92326	303.8524	0.1414	233.5950	-21.8447	-1.4122	0.7092	23.198
RandomForest	4286	65.4678	0.0422	47.5477	9.8003	0.6208	0.9865	7.319
Adaboost-PSO-BP	10056	100.2786	0.0528	81.7651	6.8082	0.4321	0.9683	37.455
Adaboost-GA-ELM	6747.6	82.1443	0.0467	63.4580	-1.9747	-0.1260	0.9787	68.177
Adaboost-PSO-ELM	1308.6	36.1751	0.0180	24.9379	0.1356	0.0864	0.9959	190.709

To comprehensively evaluate the prediction performance of the Adaboost-PSO-ELM model proposed in this article, Table 3 lists the error indicators and training time of it and several other prediction models. Compared with the optimized neural network, the error indexes of the PSO-ELM model are better than that of the PSO-BP, GPR and GA-ELM models. In terms of the coefficient of determination representing the predictive performance of the model, PSO-ELM is also the best performer, which makes it more feasible to be combined with the Adaboost algorithm.

In the ensemble learning models, since the performance of PSO-ELM is better than PSO-BP and GA-ELM, combined with the Adaboost algorithm, the performance of Adaboost-PSO-ELM is also better than Adaboost-PSO-BP and Adaboost-GA-ELM. Among the three ensemble models based on decision trees, the random forest algorithm has the best prediction performance, but it is still not as good as the Adaboost-PSO-ELM model. Fig. 13 shows the error box plot of all the prediction models above. It can be seen that the boxplot of the Adaboost-PSO-ELM model is shorter and the error distribution is closer to 0, which means that the prediction effect of this model is better than other models.

In addition, training speed is also an important indicator for evaluating the performance of wind power prediction models. From the comparison of training time in Table 3, it can be seen that the training time of the PSO-SVM model is too long, which makes it practical to be used in the combination with the Adaboost algorithm, and it is difficult to achieve real-time optimization and update in practical applications. In the optimized neural network models, the predictive performance and training time of GA-ELM and GPR are relatively close. Although the training time of PSO-BP is shorter, its error indexes are not as good as these two models. PSO-ELM achieves the best prediction performance while controlling the training time within an acceptable range, which also creates conditions for the further optimization of the Adaboost algorithm. In the ensemble learning models, although the prediction performance of random forest algorithm is not as good as that of Adaboost-PSO-ELM, its training time is much shorter, and it has broad application prospects in the ultra short term prediction of wind power.

C. LARGE SAMPLE WIND POWER PREDICTION MODEL BASED ON ADABOOST-PSO-ELM MODEL

The change of seasons will lead to changes in meteorological conditions, which will affect the output of wind turbines. The wind speed, wind direction, and wind power variation curves of the training samples corresponding to the four seasons are shown in Fig. 14. And Fig. 15 and Fig. 16 show the average wind speed and monthly sum of wind power of the wind farm every month in a year. It can be seen that the average wind

TABLE 4. Comparison of prediction error indexes in four seasons.

Model	Date	Error Indexes						
		MAE	MSE	RMSE	MAPE	MBE	RMBE	R^2
GPR	2018/2/15	19.0501	754.1664	27.4621	0.1577	-1.3083	-0.5841	0.8201
	2018/5/15	60.8211	7121.9	84.3914	0.1054	42.7978	6.2134	0.9689
	2018/8/15	59.3739	5467.4	73.9415	0.0195	26.294	0.8085	0.9577
	2018/11/15	73.3909	9221.4	96.028	0.1043	57.4302	7.2505	0.9527
PSO-BP	2018/2/15	13.3188	450.2214	21.2184	0.0903	-1.6741	-5.942	0.8925
	2018/5/15	94.5762	15676	125.2038	0.1294	83.3972	11.4337	0.9316
	2018/8/15	87.7529	14774	121.5466	0.0331	70.5451	2.3819	0.8856
	2018/11/15	101.7233	15954	126.3093	0.1241	90.7723	10.997	0.9181
PSO-SVM	2018/2/15	17.0003	574.1285	23.961	0.1277	10.4185	25.8739	0.863
	2018/5/15	101.1517	17751	133.2347	0.1318	95.1601	12.8393	0.9225
	2018/8/15	96.8646	12184	110.3813	0.0309	81.7544	2.4716	0.9056
	2018/11/15	109.0849	17996	134.1495	0.1326	101.6157	12.151	0.9076
GA-ELM	2018/2/15	13.9016	523.7749	22.8861	0.0932	-5.378	-21.9779	0.875
	2018/5/15	79.875	12528	111.9267	0.1025	58.922	8.3586	0.9453
	2018/8/15	106.5645	14597	120.8188	0.0345	90.4486	2.7273	0.8869
	2018/11/15	87.6107	12355	111.1515	0.1051	64.7797	8.1032	0.9366
PSO-ELM	2018/2/15	7.7112	204.0508	14.2846	0.0529	-0.0235	-0.0788	0.9513
	2018/5/15	35.5152	3017.1	54.9272	0.0537	5.3051	0.8212	0.9868
	2018/8/15	34.9545	2025.5	45.005	0.0107	-0.137	0.4257	0.9843
	2018/11/15	30.4978	2218.6	47.1015	0.038	18.6901	-2.5441	0.9886
Bagged Trees	2018/2/15	16.168	706.5731	26.5814	0.1309	11.7046	28.1682	0.8313
	2018/5/15	40.7745	3109.4	55.7624	0.079	34.2186	5.0305	0.9864
	2018/8/15	45.2984	3329.3	57.7003	0.0145	2.8907	0.0895	0.9742
	2018/11/15	39.1085	2897	53.8242	0.0527	32.0396	4.1789	0.9851
Boosting Trees	2018/2/15	42.2657	2916.5	54.0042	0.3049	15.766	34.5639	0.3038
	2018/5/15	169.0757	60844	246.6651	0.2292	52.9679	7.578	0.7345
	2018/8/15	205.9284	59728	244.3926	0.0676	-65.6967	-2.0788	0.5374
	2018/11/15	554.8935	50697	712.0214	0.6461	-441.6774	-150.7529	0.4761
Random Forest	2018/2/15	15.2474	616.7066	24.8336	0.1213	11.2135	27.3089	0.8528
	2018/5/15	38.7593	2964.9	54.4508	0.0704	31.7676	4.6871	0.9871
	2018/8/15	43.4201	3008.8	54.8523	0.0138	9.3648	0.2895	0.9767
	2018/11/15	38.9169	2855	53.4322	0.0526	31.6311	4.1278	0.9853
Adaboost-PSO-BP	2018/2/15	15.8235	522.3914	22.8559	0.1151	-8.1647	-37.6545	0.8753
	2018/5/15	82.4056	12998	114.0076	0.1022	66.3376	9.3127	0.9433
	2018/8/15	102.4173	10584	102.536	0.0282	78.7127	2.1401	0.9186
	2018/11/15	79.3725	11052	105.1307	0.0917	62.9731	7.895	0.9433
Adaboost-GA-ELM	2018/2/15	10.2392	377.4813	19.4289	0.0636	0.8675	2.8244	0.9099
	2018/5/15	36.5476	3233.4	56.8633	0.0593	0.956	0.1478	0.9859
	2018/8/15	43.6358	3081.8	55.5143	0.0141	1.227	0.038	0.9761
	2018/11/15	34.9721	2652.6	51.5035	0.0545	0.8945	0.1216	0.9864
Adaboost-PSO-ELM	2018/2/15	6.2742	109.862	10.4815	0.0468	-0.00239	-0.0801	0.9738
	2018/5/15	21.7115	1037.5	32.2103	0.031	0.0111	-0.0715	0.9955
	2018/8/15	28.0692	1354.1	36.7973	0.0086	-0.0981	0.0304	0.9895
	2018/11/15	23.3811	1023.4	31.9903	0.0286	-0.3222	0.04387	0.9947

speed in summer is the lowest, and the corresponding wind power is also the lowest. Due to the cut-in speed of the wind turbine, the turbine will only output power when the wind speed is greater than the cut-in speed. Therefore, there is a situation where the wind speed is not equal to 0 but the turbine's output power is 0, which brings unfavorable factors to the prediction.

To verify the applicability to samples of different seasons of the proposed method, in the historical data of February, May, August, and November, 3.5 consecutive days of data are selected each month, and a total of 2016 sets of data are selected to be training samples. The short term wind power data on the 15th day of the four months was used as test sample. It is verified in Section B that the Adaboost-PSO-ELM wind power prediction model has higher prediction accuracy and wider applicability than other optimized neural networks and ensemble learning models in a specific set of samples. It remains to be proved whether the Adaboost-PSO-ELM model can maintain its superiority in the prediction of a larger sample.

The histograms of monthly average wind speed and monthly wind power in Fig. 15 and Fig. 16 show that the output power of wind turbines is closely related to the level of local wind speed. The average wind speed from April to July is relatively low, and the corresponding monthly wind power is also low. It can also be seen that the higher wind power occurs in the months with higher average wind speed. The comparison of the wind power prediction results of the Adaboost-PSO-ELM model and several current prediction models in different seasons is shown in Fig. 17 and Fig. 18. To quantify and evaluate the prediction performance of various models, the analysis of the error indexes when predicting the four season samples in Table 4 is as follows:

(1) The rated wind speed is the wind speed when the wind turbine reaches its rated power. When processing the samples in May and November, the wind direction is relatively stable,

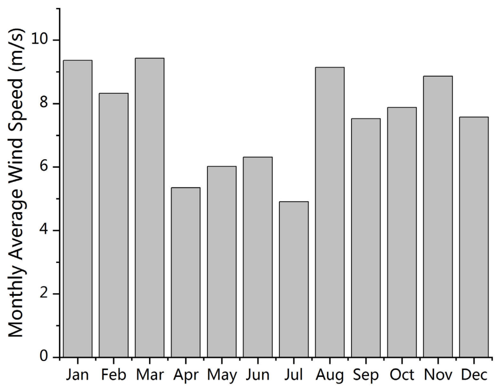

FIGURE 15. Average wind speed in twelve months.

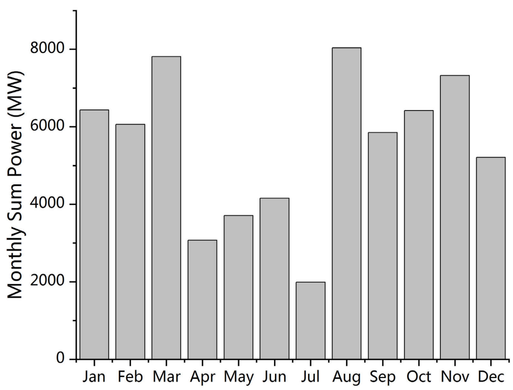

FIGURE 16. Monthly sum of wind power in twelve months.

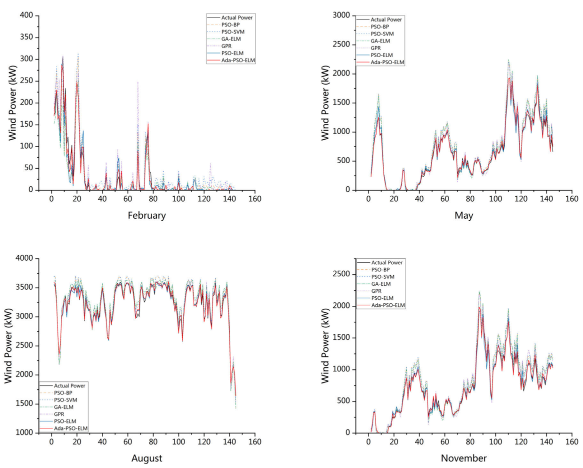

FIGURE 17. Comparison of neural networks' prediction in four seasons.

and the wind speed is between the cut-in wind speed and the rated wind speed for most of the time, which is conducive to the function of the regression model. Therefore, most models performs relatively well in these two months.

(2) The coefficient of determination R^2 is an intuitive manifestation of the predictive performance of the model. From the comparison of R^2, it can be seen that the Adaboost-PSO-ELM model maintains a better prediction performance than other

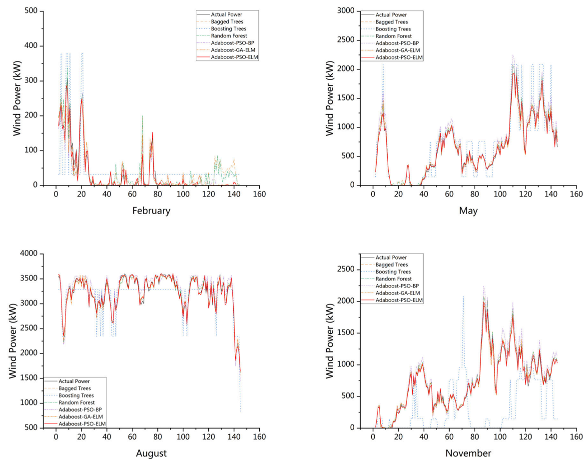

FIGURE 18. Comparison of ensemble learning models' prediction in four seasons.

models in the prediction of the four season samples. Due to the fact that there are many sample points with a power value of 0 in February, which has an adverse impact on the prediction, the coefficient of determination of each model for the February sample prediction is generally 2%-15% lower than that of other months. When predicting the other three months, the coefficients of determination of each model are relatively balanced. But in August (autumn), it can be seen from Fig. 14 that the wind direction changes frequently and with a large magnitude, resulting in a lower coefficient of determination for August than in May and November. This influence on R^2 is manifested in each model's prediction results of August.

(3) Mean absolute error (MAE) is the average of the absolute value of the error between the actual and predicted values of each sample. It represents the modulus length of the prediction error, but does not consider the direction of the error. If the absolute value sign is removed, the index becomes the mean bias error (MBE), which represents the dispersion between the predicted value and the actual value. It can be seen from Table 4 that the MAE and MBE value of most models are similar, but the MBE value of Adaboost-GA-ELM and Adaboost-PSO-ELM is much smaller than their MAE value, which indicates the fluctuation of their prediction results is closer to reality.

(4) The mean square error (MSE) represents the sum of the squares of the distance between the predicted value and the actual value, and the root mean square error (RMSE) is the square root of the MSE. Since the power and average wind speed in February are the smallest of these four months, the MSE and RMSE values of the models for the February sample are also the smallest. Relative mean bias error (RMBE) is the ratio of the prediction error of each sample to the actual value. The smaller the value represents the higher prediction accuracy. In the comparison of these three indicators, Adaboost-PSO-ELM is the best performer.

(5) The mean absolute percentage error (MAPE) is expressed as a percentage and not relative to the value of

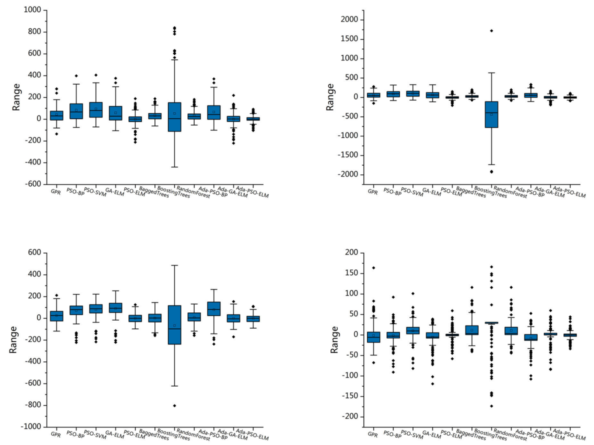

FIGURE 19. Error box diagram of prediction in four seasons.

the samples. It can be used to compare the prediction performance of samples with very different values. But MAPE is asymmetric, it is more sensitive to negative errors (when the predicted value is higher than the actual value) than to positive errors. This is also the reason why the MAPE and R^2 value are both low in each model's prediction of the sample in August. Since there are sample points with a power of 0 in the samples of February, May, and November, the MAPE value is meaningless. The MAPE value is calculated after the data points with a value of 0 are deleted. In the comprehensive performance of this index, Adaboost-PSO-ELM model is also better than other optimized neural networks and ensemble learning models.

Through the comparison of the several error indexes above and the error box diagram in Fig. 19, it can be seen that the Adaboost-PSO-ELM model is suitable for wind power prediction under various conditions and has better prediction performance than other models. This shows that the Adaboost-PSO-ELM model can better adapt to sudden changes in wind speed and direction, and has good stability and generalization ability.

IV. CONCLUSION

Aiming at the volatility of wind power and the poor prediction accuracy and generalization ability of existing wind power prediction models, this paper proposes a wind power prediction method based on Adaboost-PSO-ELM integrated learning model. The following conclusions are obtained through the analysis of the measured wind power data:

(1) Using a small sample data set for training, the Adaboost-PSO-ELM wind power prediction model performs better than GPR, PSO-BP, PSO-SVM, GA-ELM, PSO-ELM, Boosting Trees, Bagged Trees, Random Forest, Adaboost-PSO-BP and Adaboost-GA-ELM in MAE, MSE, RMSE, MAPE, MBE, RMBE and R^2 indexes.

(2) When training on a large sample data set containing multi-season data, the Adaboost-PSO-ELM wind power prediction model can keep up with the changes in wind speed and direction in each season and make predictions, which is superior to other traditional models. This model's prediction performance proves that it has good robustness and generalization ability and can provide a more reliable basis for power grid dispatch.

(3) In the current wind power prediction models, the training samples are still mainly selected based on experience. In future work, we plan to consider reconstructing the training samples. Based on numerical weather forecast (NWP) data, similar days are selected according to specific indicators as training samples for short-term wind power prediction to further improve prediction performance.

REFERENCES

[1] X. Yuan, C. Chen, Y. Yuan, Y. Huang, and Q. Tan, "Short-term wind power prediction based on LSSVM–GSA model," *Energy Convers. Manage.*, vol. 101, pp. 393–401, Sep. 2015.

[2] *2020 Global Wind Report*, Global Wind Energy Council (GWEC), Brussels, Belgium, 2021, pp. 6–7.

[3] Y. Xue, Y. Chen, J. Zhao, X. Liu, K. Li, W. Qiu, and Y. Gang, "A review on short-term and ultra-short-term wind power prediction," *Autom. Electr. Power Syst.*, vol. 39, no. 6, pp. 141–151, 2015.

[4] B. Zhou, X. Ma, Y. Luo, and D. Yang, "Wind power prediction based on LSTM networks and nonparametric kernel density estimation," *IEEE Access*, vol. 7, pp. 165279–165292, 2019.

[5] H. Yan and Z. Wu, "A hybrid short-term wind power prediction model combining data processing, multiple parameters optimization and multi-intelligent models apportion strategy," *IEEE Access* vol. 8, pp. 227126–227140, 2020.

[6] J. Lee, W. Wang, F. Harrou, and Y. Sun, "Wind power prediction using ensemble learning-based models," *IEEE Access*, vol. 8, pp. 61517–61527, 2020.

[7] Z. Sun and M. Zhao, "Short-term wind power forecasting based on VMD decomposition, ConvLSTM networks and error analysis," *IEEE Access*, vol. 8, pp. 134422–134434, 2020.

[8] J. Shi, Y. Liu, Y. Yang, and W. Lee, "Short-term wind power prediction based on wavelet transform-support vector machine and statistic characteristics analysis," in *Proc. IEEE Ind. Commercial Power Syst. Tech. Conf.*, May 2011, pp. 1–7.

[9] T. Ouyang, X. Zha, L. Qin, Y. He, and Z. Tang, "Prediction of wind power ramp events based on residual correction," *Renew. Energy*, vol. 136, pp. 781–792, Jun. 2019.

[10] A. Tascikaraoglu and M. Uzunoglu, "A review of combined approaches for prediction of short-term wind speed and power," *Renew. Sustain. Energy Rev.*, vol. 34, pp. 243–254, Jun. 2014.

[11] J. Wang, J. Hu, K. Ma, and Y. Zhang, "A self-adaptive hybrid approach for wind speed forecasting," *Renew. Energy*, vol. 78, pp. 374–385, Jun. 2015.

[12] E. T. Renani, M. F. M. Elias, and N. A. Rahim, "Using data-driven approach for wind power prediction: A comparative study," *Energy Convers. Manage.*, vol. 118, pp. 193–203, Jun. 2016.

[13] Y. Zhang, H. Sun, and Y. Guo, "Wind power prediction based on PSO-SVR and grey combination model," *IEEE Access*, vol. 7, pp. 136254–136267, 2019.

[14] L. Tan, J. Han, and H. Zhang, "Ultra-short-term wind power prediction by salp swarm algorithm-based optimizing extreme learning machine," *IEEE Access*, vol. 8, pp. 44470–44484, 2020.

[15] N. Lu and Y. Liu, "Application of support vector machine model in wind power prediction based on particle swarm optimization," *Discrete Continuous Dyn. Syst. S*, vol. 8, no. 6, pp. 1267–1276, 2015.

[16] J. Li and J. Mao, "Ultra-short-term wind power prediction using BP neural network," in *Proc. IEEE Conf. Ind. Electron. Appl.*, Jun. 2014, pp. 2001–2006.

[17] S. Wang, N. Zhang, L. Wu, and Y. Wang, "Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method," *Renew. Energy*, vol. 94, pp. 629–636, Aug. 2016.

[18] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.

[19] X. Wang, C. Wang, and Q. Li, "Short-term wind power prediction using GA-ELM," *Open Electr. Electron. Eng. J.*, vol. 11, no. 1, pp. 48–56, Feb. 2017.
[20] X. Zhai and L. Ma, "Medium and long-term wind power prediction based on artificial fish swarm algorithm combined with extreme learning machine," *Int. Core J. Eng.*, vol. 5, no. 10, pp. 265–272, 2019.
[21] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, pp. 119–139, Aug. 1995.
[22] X. Li and Q.-Y. Zhu, "Tax forecasting based on AdaBoost algorithm and BP neural network," *J. Comput. Appl.*, vol. 32, no. 12, pp. 3558–3560, May 2013.
[23] T. Wen and F. Yu, "Research on prediction of water inrush from coal seam floor based on PSO_SVM_AdaBoost," *Appl. Res. Comput.*, vol. 35, no. 12, pp. 3664–3667 and 3677, Dec. 2017.
[24] G.-B. Huang, D. H. Wang, and Y. Lan, "Extreme learning machines: A survey," *Int. J. Mach. Learn. Cybern.*, vol. 2, no. 2, pp. 107–122, Jun. 2011.
[25] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. 6th Int. Symp. Micro Mach. Hum. Sci. (MHS)*, 2002, pp. 39–43.
[26] H. Liu, Z. Duan, Y. Li, and H. Lu, "A novel ensemble model of different mother wavelets for wind speed multi-step forecasting," *Appl. Energy*, vol. 228, pp. 1783–1800, Oct. 2018.
[27] Z. Yang and Y. Liu, "Short-term wind power prediction with particle swarm optimization," *Power Syst. Technol.*, vol. 35, no. 5, pp. 159–164, May 2011.
[28] H. Wang, Z. Hu, M. Hu, and Z. Zhang, "Short-term prediction of wind farm power based on PSO-SVM," in *Proc. Asia–Pacific Power Energy Eng. Conf.*, Mar. 2012, pp. 1–4.
[29] Z. Ye and Y. Ding, "Short-term wind power prediction based on nutrosophic clustering and GA-ELM," *J. Phys., Conf. Ser.*, vol. 1607, Aug. 2020, Art. no. 012025.

GUOQING AN was born in 1982. He received the B.S. degree in electrical machines and apparatus from the Hebei University of Science and Technology, Shijiazhuang, in 2007, and the Ph.D. degree from the Hebei University of Technology, Tianjin, in 2013.

He joined the Hebei University of Science and Technology, in 2007, as an Assistant, where he was promoted to a Lecturer, in 2009, and he has been an Associate Professor with the College of Electrical Engineering, since 2015. His research interests include wind power prediction and the condition monitoring of power equipment.

ZIYAO JIANG was born in 1996. He is currently pursuing the M.S. degree with the Hebei University of Science and Technology. His main research interests include wind power prediction and the application of machine learning in wind energy.

XIN CAO received the Ph.D. degree in national economics from the Renmin University of China, Beijing, China, in 2003.

He has been a Senior Economist with Hebei Construction & Investment Group New Energy Company Ltd., Shiajiazhuang, China, and Chongli Xintian Wind Energy Corporation Ltd., Shijiazhuang. He is currently the Party Secretary, the President, and the General Manager with Hebei Construction & Investment Group New Energy Company Ltd. His major research interests include wind power generation and hydrogen production.

YUFEI LIANG is currently pursuing the M.S. degree with the Hebei University of Science and Technology. His main research interests include non-intrusive load monitoring technology and wind power prediction.

YUYANG ZHAO received the M.S. and Ph.D. degrees in control science and engineering from Tianjin University, China. She has been undertaking research with the University of Huddersfield, as a Visiting Scholar, which was sponsored by the China Scholarship Council. She is currently a Lecturer with the Hebei University of Science and Technology. Her main research interests include electrical engineering and control science and technology.

ZHENG LI (Member, IEEE) was born in Shijiazhuang, Hebei, China, in 1980. He received the B.Sc. and Ph.D. degrees in electrical engineering and power electronics and electric drive from the Hefei University of Technology, Hefei, China, in 2002 and 2007, respectively.

Since 2007, he has been a Lecturer, an Associate Professor, and a Professor with the School of Electrical Engineering, Hebei University of Science and Technology. From July 2013 to July 2014, he has been a Visiting Scholar and a part-time Faculty with the College of Engineering, Wayne State University, USA. He is the author of more than 240 published articles. His current research interests include design, analysis, and control of novel motors and actuators, intelligent control, and power electronics.

Dr. Li is an Active Reviewer of the IEEE Transactions on Industrial Electronics, IEEE Transactions on Energy Conversion, and IEEE Transactions on Magnetics and Electric Power Components and Systems.

WEICHAO DONG was born in Shijiazhuang, Hebei, China, in 1989. He received the B.Sc. degree in electrical engineering and the M.Sc. degree in electrical and computing engineering from Cornell University, NY, USA, in 2012 and 2013, respectively. He is currently pursuing the Ph.D. degree in control theory and control engineering with the Hebei University of Technology, Tianjin, China. He is also a Lecturer with the College of Electrical Engineering, Hebei University of Science and Technology, China. He has published two articles in SCI. His research interests include the optimization of wind engine structure, wind power, and the application of deep learning in the wind energy.

HEXU SUN (Senior Member, IEEE) received the Ph.D. degree in automation from Northeastern University, China, in 1993. He is currently a Professor with the School of Artificial Intelligence, Hebei University of Technology. He is also a Professor with the Hebei University of Science and Technology. His current research interests include wind energy and complex engineering systems.