

Received November 25, 2020, accepted December 4, 2020, date of publication December 10, 2020, date of current version December 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3043812

A Hybrid Model Based on Multi-Stage Principal Component Extraction, GRU Network and KELM for Multi-Step Short-Term Wind Speed Forecasting

FENG ZOU^{1,2}, WENLONG FU^{1,2}, (Member, IEEE), PING FANG^{1,2},
DONGZHEN XIONG^{1,2}, AND RENMING WANG¹

¹College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China

²Hubei Provincial Key Laboratory for Operation and Control of Cascaded Hydropower Station, China Three Gorges University, Yichang 443002, China

Corresponding author: Wenlong Fu (ctgu_fuwenlong@126.com)

This work was supported by the National Natural Science Foundation of China under Grant 51741907.

ABSTRACT Accurate wind speed forecasting exerts a critical role in energy conversion and management of wind power. In term of this purpose, a hybrid model based on multi-stage principal component extraction, kernel extreme learning machine (KELM) and gated recurrent unit (GRU) network is developed in this paper, where the multi-stage principal component extraction combines complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), singular spectrum analysis (SSA) and phase space reconstruction (PSR). Firstly, CEEMDAN is employed to decompose the raw wind speed data into a sequence of intrinsic mode functions (IMFs) and a residual component. Then the principal components and residual components of all IMFs are captured by SSA. Meanwhile, all residual components obtained by CEEMDAN decomposition and SSA processing are added to form a new component. Subsequently, PSR is utilized to construct each forecasting component obtained by CEEMDAN-SSA into the input and output of training set and testing set for the prediction model. Later, KELM and GRU neural network are conducted to predict the high-frequency and low-frequency components, respectively. Eventually, the prediction values of each component are accumulated to acquire the final prediction result. To evaluate the performance of the proposed model, four datasets from Sotavento Galicia wind farm are adopted to conduct experimental research. The experimental results manifest that the proposed model achieves higher accuracy of multi-step prediction than other comparative models.


INDEX TERMS Multi-step short-term wind speed prediction, multi-stage principal component extraction, complete ensemble empirical mode decomposition with adaptive noise, singular spectrum analysis, phase space reconstruction, kernel extreme learning machine, gated recurrent unit network.

I. INTRODUCTION

With the express development of the world economy, great changes have taken place in the energy structure. Meanwhile, the environmental problems and energy crisis caused by the massive consumption of traditional fossil energy have attracted extensive attention [1]. Therefore, finding renewable energy that can replace fossil energy has become the focus of current research. With the green and pollution-free characteristics, wind energy has turned into one of the most rapidly developing renewable energy sources [2]. However, due to the intermittency and randomness of wind speed, wind

power grid connection produces a negative impact on the safe and stable operation of the power system [3]. Through relevant studies, accurate wind speed prediction can decrease the negative influence brought by wind power grid connection, which is of great significance to maintain the secure and steady operation of the power system [4].

In recent years, a large number of researches have been carried out by many scholars in the field of wind speed forecasting [5]. Generally, these prediction approaches can be divided into four categories: (1) physical models; (2) statistical models; (3) artificial intelligence models; (4) hybrid models. Physical models are modelled by physical formulas based on relevant parameters [6], such as temperature, humidity, pressure and topography, etc., among which one

The associate editor coordinating the review of this manuscript and approving it for publication was Jinming Wen .

of the most widely applied models is numerical weather prediction (NWP) [7]. However, the physical models possess disadvantages in short-term wind speed forecasting because of their complexity and professionalism and need to consume many computing resources [8]. By contrast, statistical models utilize historical wind speed data to predict current wind speed data, which is easy to implement and more suitable for short-term wind speed forecasting [9], [10]. The traditional statistical models mainly include autoregressive (AR) [11], autoregressive moving average (ARMA) [12], autoregressive integrated moving average (ARIMA) [13] and fractional autoregressive integrated moving average (FARIMA) [14] and they are utilized for wind speed prediction research by many scholars in recent years. Specifically speaking, Maatalah *et al.* [15] combined the Hammerstein model with AR model to propose a new wind speed prediction model, and finally obtained a higher prediction accuracy. Two hybrid models combining ARMA model with the non-parametric model were presented by Han *et al.* [16], and the experimental results demonstrated that the prediction results of the hybrid model are better than that of comparing models. Nevertheless, the statistical model is commonly employed for the prediction of linear time series, and it is not appropriate for directly forecasting wind speed with nonlinear as well as nonstationary characteristics [17]. At present, artificial intelligence (AI) technology develops rapidly and has been widely applied in many fields and achieves fine effects. These models mainly include back propagation neural network (BPNN) [18], support vector machine (SVM) [19], [20], extreme learning machine (ELM) [21], Gaussian Process Regression (GPR) [22] as well as long short-term memory network (LSTM) [23], and many of which are applied for studying wind speed forecasting. To be specific, Zhou *et al.* [24] developed a systematic research to regulate the parameters of SVM in short-term wind speed prediction, and results showed a better forecasting performance. Guo *et al.* [18] proposed a hybrid method utilizing BPNN which improved the forecasting accuracy of wind speed effectively. A novel prediction method based on ELM was proposed by Fu *et al.* [3]. The experimental results demonstrated that the proposed method can achieve higher forecasting accuracy in multi-step wind speed prediction. Although the AI method has better predictive ability than the linear model, it still has some problems, such as easy to fall into local optimal or overfitting, low convergence speed, and some key parameters that are difficult to determine [25], [26].

According to relevant research, each model has its own advantages and disadvantages, and it is hard to acquire ideal experimental results by directly utilizing them for prediction. In order to further improve the prediction accuracy of wind speed time series, data preprocessing techniques and optimization algorithms are applied to wind speed prediction methods to form hybrid prediction models [27], [28], [29]. Among these data preprocessing techniques, decomposition algorithms are commonly employed, such as wavelet decomposition (WD) [30], empirical mode decomposition

(EMD) [31], wavelet packet decomposition (WPD) [32] and ensemble empirical mode decomposition (EEMD) [33] etc. In addition, the optimization algorithms mainly include particle swarm optimization (PSO) [34], genetic algorithm (GA) [35], harris hawks optimization (HHO) [36] etc. Based on this, several hybrid models have been proposed. To be specific, Wang *et al.* [33] presented a hybrid prediction model including EEMD and GA for wind speed prediction, and the experimental results revealed that the presented model can effectively boost the prediction performance. Two novel hybrid models proposed by Zhang *et al.* [37] for short-term wind speed forecasting was composed of feature selection, EMD, ANN and SVM. It can be concluded that the proposed models achieved satisfactory prediction accuracy from the experimental results, which is suitable for short-term wind speed prediction. Moreover, some hybrid models are proved to have superior predictive performance by combining singular spectrum analysis (SSA). For instance, a hybrid model composed of EMD and SSA was developed by Yu *et al.* [25], in which SSA was adopted to process the highest frequency component decomposed. The experimental results indicated that the proposed model can achieve higher prediction accuracy. Yu *et al.* [38] developed a hybrid model including wavelet transform (WT) and SSA, where SSA was implemented to capture dominant components of the highest frequency subseries decomposed by WT and the experiments results demonstrated that the performance of the hybrid model is better than that of other models. Although the aforementioned EMD and EEMD decomposition algorithms can promote the forecasting accuracy of the hybrid model to a certain degree, they still exist some shortcomings which can be summarized as the mode mixing problem of EMD as well as the residual noise problem of EEMD. Moreover, some scholars have presented a variety of hybrid models based on dual predictors and obtained good experimental effect. For example, Liu *et al.* [39] successfully designed a hybrid multi-step forecasting model using LSTM network and ELM as predictors, where LSTM network was adopted to predict the low-frequency sub-components acquired by decomposing while ELM was conducted to predict the high-frequency sub-components acquired by decomposing. The experimental results indicated that the designed model achieved a satisfactory predictive effect. A hybrid model using convolutional long short-term memory (ConvLSTM) network as well as kernel extreme learning machine (KELM) as predictors was developed by Fu *et al.* [40], where KELM was implemented to predict high-frequency components while ConvLSTM was employed to predict low-frequency components. The experimental results demonstrated that the hybrid model possesses higher prediction accuracy.

Inspired by the above analysis, the hybrid model combining complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), SSA, phase space reconstruction (PSR), KELM and gated recurrent unit (GRU) neural network is proposed in this paper. First of all, CEEMDAN and SSA are utilized for multi-stage principal component

extraction of the raw wind speed data, in which CEEMDAN is adopted to decompose the raw wind speed data into a sequence of intrinsic mode functions (IMFs) and a residual component while SSA is adopted to further capture principal components and residual components of each IMFs, and all the residual components obtained by CEEMDAN decomposition and SSA processing are added to form a new predicted component. Meanwhile, among all the new IMFs components, the top five components with the biggest fluctuation are taken as high-frequency components, while the rest are treated as low-frequency ones. Secondly, KELM and GRU neural network are utilized to forecast the high-frequency and low-frequency components reconstructed by PSR, respectively. Ultimately, the forecasting results of all components are superimposed to acquire the final predicted values. The main contributions of this study can be summarized as follows:

1) Considering the nonlinear and non-stationarity of raw wind speed data, the multi-stage principal component extraction method based on CEEMDAN and SSA is utilized to promote the predictive capability of proposed hybrid model for multi-step wind speed forecasting effectively.

2) PSR is implemented to reconstruct the wind speed data with chaotic characteristics into the input and output of training set and testing set for the proposed model.

3) KELM with high predictive efficiency is conducted to forecast the high-frequency components acquired by CEEMDAN-SSA while GRU neural network with strong nonlinear processing capability is conducted to forecast the low-frequency components acquired by CEEMDAN-SSA.

The remaining of this paper is arranged as follows: Section II introduces relevant methodology including CEEMDAN, SSA, PSR, KELM and GRU neural network. Section III presents the detailed procedures of hybrid multi-step forecasting model in this study. The case study section discusses experiment results and comparative analysis of the proposed hybrid method for wind speed prediction. Finally, the conclusion of this paper is summarized in the last section.

II. METHODOLOGY

In this section, the theoretical methods related to the proposed hybrid prediction model will be introduced, including CEEMDAN, SSA, PSR, KELM and GRU neural network.

A. COMPLETE ENSEMBLE EMPIRICAL MODE DECOMPOSITION WITH ADAPTIVE NOISE

CEEMDAN is a signal processing method proposed by Torres *et al.* [41] which decomposes the original signal into intrinsic mode functions (IMFs). By adding a finite amount of adaptive white noise to each decomposition process, this method not only overcomes the problem of model mixing caused by EMD decomposition but also addresses the problem of incompleteness caused by EEMD decomposition as well as low computational efficiency created by increasing the average number of times. It can be used to

analyze nonlinear and nonstationary time series with high signal-to-noise ratio and owns good time-frequency focusing.

Set $x(t)$ as the original signal, which is added with Gaussian white noise $\varepsilon_0\omega(t)^{(i)}$ obeying standard normal distribution. Then the expression of the i -th sequence can be depicted as $x(t)^{(i)} = x(t) + \varepsilon_0\omega(t)^{(i)}$, ($i = 1, \dots, I$), where $\omega(t)^{(i)}$ represents a set of Gaussian white noise sequences with zero mean and unit variance, and ε_0 denotes the standard deviation of Gaussian white noise. The specific steps can be formulated as follows:

Step 1: The EMD decomposition is performed for $x(t)^{(i)}$ to obtain the first IMF and take its mean value as the first IMF₁.

$$\text{IMF}_1 = \frac{1}{I} \sum_{i=1}^I E_1[x(t)^{(i)}] \quad (1)$$

$$r_1 = x(t) - \text{IMF}_1 \quad (2)$$

where E is the EMD decomposition operator, r_1 denotes the residual signal after the first decomposition.

Step 2: Adding specific noise to the new signal, EMD decomposition is continued to obtain the second IMF₂ of the original signal and corresponding residual signal r_2 .

$$\text{IMF}_2 = \frac{1}{I} \sum_{i=1}^I E_1(r_1 + \varepsilon_1 E_1[\omega(t)^{(i)}]) \quad (3)$$

$$r_2 = r_1 - \text{IMF}_2 \quad (4)$$

Step 3: Similarly, for $i = 3, \dots, k$, calculate the k -th mode component and corresponding residual signal r_k according to step 2.

$$\text{IMF}_k = \frac{1}{I} \sum_{i=1}^I E_1(r_{k-1} + \varepsilon_{k-1} E_{k-1}[\omega(t)^{(i)}]) \quad (5)$$

$$r_k = r_{k-1} - \text{IMF}_k \quad (6)$$

Step 4: Repeat step 3 until the residual signal no more meets the EMD decomposition condition that the number of IMFs local extremum points is less than 3, and the algorithm is terminated. Hence, the decomposition consequence of the original signal can be described as:

$$x(t) = \sum_{k=1}^K \text{IMF}_k + r_k \quad (7)$$

B. SINGULAR SPECTRUM ANALYSIS

SSA algorithm, a powerful method for studying nonlinear time series data, can be widely employed to solve many varieties of problems, including principal or pseudo-periodic component detection and capture, signal denoising, forecasting and edge point detection [30], [42]. Because of its strong processing ability to nonlinear signal, many scholars have applied it to the field of wind speed prediction [40]. The standard SSA includes four steps: embedding, singular value decomposition (SVD), grouping and diagonal averaging and its specific steps can be described as follows:

1) embedding. Assume that the time series of wind speed with sample number N is set as $Y = [y_1, y_2, \dots, y_N]^T$ and

the embedded dimension of SSA is represented by L , where $2 \leq L \leq N$. Then the lag order L vector can be defined as $X_i = [y_i, y_{i+1}, \dots, y_{i+L-1}]^T$, where $i = 1, 2, \dots, K$, $K = N - L + 1$. Thus, the raw wind speed data is reconstructed into a trajectory matrix named Hankel matrix, which can be defined as follows:

$$X = \begin{pmatrix} y_1 & y_2 & \cdots & y_K \\ y_2 & y_3 & \cdots & y_{K+1} \\ \vdots & \vdots & \dots & \vdots \\ y_L & y_{L+1} & \cdots & y_N \end{pmatrix} \quad (8)$$

2) SVD. Through SVD, the Hankel matrix X can be decomposed into d components, where $d = \text{rank}(X)$. Additionally, the i -th triplet eigenvector of the SVD for the matrix X is (λ_i, U_i, V_i) , where λ_i denotes the i -th eigenvalue of the covariance matrix XX^T , U_i and V_i represent i -th left eigenvector as well as right eigenvector of matrix XX^T . Therefore, the Hankel matrix X can be further expressed as follows:

$$X = \sum_{i=1}^d X_i \quad (9)$$

$$X_i = \sqrt{\lambda_i} U_i V_i \quad (10)$$

3) Grouping. The d components generated by SVD are divided into main m components as the dominant components. For $I = \{I_1, I_2, \dots, I_m\}$, the matrix X_I corresponding to the group I can be defined as follows:

$$X_I = X_{I_1} + X_{I_2} + \dots + X_{I_m} \quad (11)$$

4) Diagonal averaging. In this process, each matrix grouped in Eq.(11) is transferred into the new time series by the following procedure. Assuming that X is a $P \times J$ matrix with elements x_{ij} , $P^* = \min(P, J)$, $J^* = \max(P, J)$. If $P < J$, let $x_{ij}^* = x_{ij}$, otherwise, let $x_{ij}^* = x_{ji}$. Then the reconstructed time series $Z = \{z_1, z_2, \dots, z_N\}$ can be expressed as follows:

$$z_k = \begin{cases} \frac{1}{k} \sum_{q=1}^k x_{q,k-q+1}^* & 1 \leq k \leq P^* \\ \frac{1}{P^*} \sum_{q=1}^{P^*} x_{q,k-q+1}^* & P^* \leq k \leq J \\ \frac{1}{N-k+1} \sum_{q=k-J^*+1}^{N-J^*+1} x_{q,k-q+1}^* & J^* \leq k \leq N \end{cases} \quad (12)$$

C. PHASE SPACE RECONSTRUCTION

PSR algorithm, a practical method to study chaotic systems, was proposed by Packard et al [43]. The main purpose of PSR is to reconstruct a phase space based on the sequence data produced by original dynamical system. Due to that the evolution information would be implied in the data developing process, PSR can reflect the inherent law of the original dynamical system. Thus, preprocessing wind speed with PSR by selecting the appropriate delay time τ as well as embedding dimension d , the wind speed prediction accuracy would

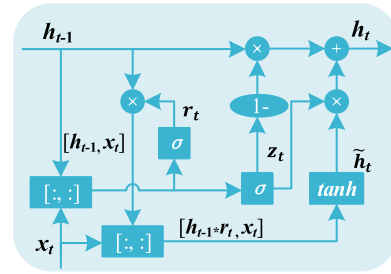


FIGURE 1. The structure of GRU unit.

be promoted. Let a one-dimensional wind speed data with length N as $x = \{x_i | i = 1, 2, \dots, N\}$, the input matrix form of PSR at different prediction levels h is expressed as follows:

$$X = [X_1 X_2 \cdots X_L]^T = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(d-1)\tau} \\ \vdots & \vdots & \ddots & \vdots \\ x_i & x_{i+\tau} & \cdots & x_{i+(d-1)\tau} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+\tau} & \cdots & x_{L+(d-1)\tau} \end{bmatrix} \quad (13)$$

where $L = N - (d - 1)\tau - h$, τ and d denote delay time and embedded dimension, respectively. X_i denotes the i -th input sample vector in the input matrix. Hence, the matched reconstructed output matrix form can be expressed as follows:

$$Y = [Y_1 Y_2 \cdots Y_L]^T = [x_{1+h+(d-1)\tau}, x_{2+h+(d-1)\tau} \cdots x_N]^T \quad (14)$$

where Y_i is the prediction value according to the i -th vector in the input matrix.

D. KERNEL EXTREME LEARNING MACHINE

ELM, a machine learning theory, was proposed by Huang et al. [44]. Based on this theory, some relevant algorithms like a modified version combined with kernel function have been derived, namely KELM. It utilizes kernel function to replace random mapping in ELM, which improves the stability and generalization ability of the model. Because the kernel function is employed, the number of hidden nodes do not need to be given, and the specific process can be shown as follow:

For the ELM model with L hidden nodes, the training samples of N groups are given as $\{(x_i, y_i) | x_i \in R^g, y_i \in R^c, i = 1, 2, \dots, N\}$. Therefore, the weight matrix β between hidden layer and the output layer can be formulated as follows:

$$\beta = H^T \left(\frac{I}{C} + HH^T \right)^{-1} T \quad (15)$$

where H is the hidden layer output matrix; T is the output target matrix; I is the identity matrix, and C is the regularization coefficient.

On the basis of ELM, the KELM method introduces the kernel function to achieve better stability and generalization performance, which can map all input samples from n -dimensional space to high-dimensional hidden layer feature

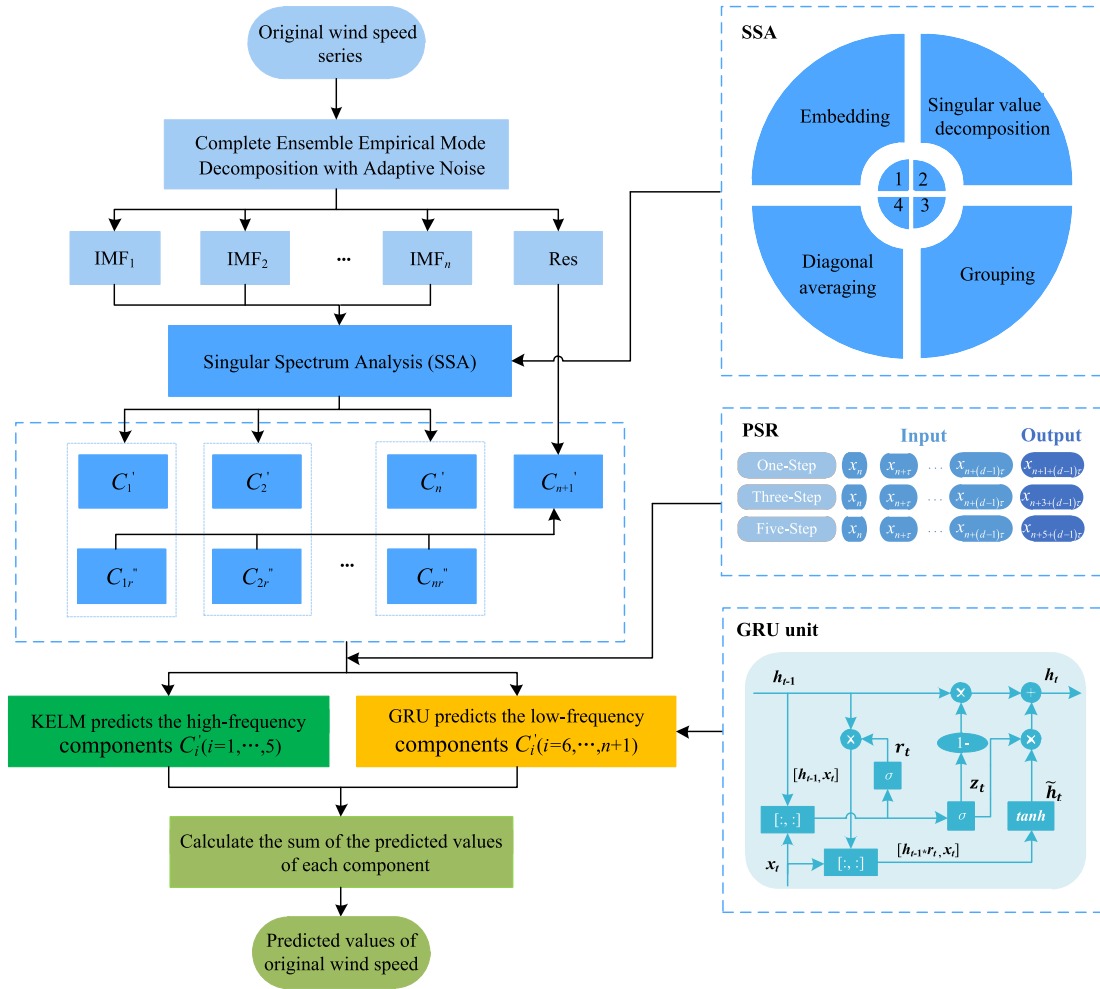


FIGURE 2. The framework of the hybrid multi-step prediction model.

space. The kernel matrix is defined as Ω_{ELM} whose element is $\Omega_{ELM}(i, j)$.

$$\Omega_{ELM} = \mathbf{H}\mathbf{H}^T; \Omega_{ELM}(i, j) = h(x_i) \cdot h(x_j) = K(x_i, x_j) \quad (16)$$

where $K(x_i, x_j)$ is the kernel function. Radial basis function is usually chosen as the kernel function of the KELM model, and its expression can be expressed as follows:

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{\delta^2}\right) \quad (17)$$

where δ represents the kernel parameter. Therefore, the ultimate output of KELM can be signified in the following:

$$\begin{aligned} f(x) &= h(x)\beta = h(x)\mathbf{H}^T \left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{T} \\ &= \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \left(\frac{\mathbf{I}}{C} + \Omega_{ELM} \right)^{-1} \mathbf{T} \end{aligned} \quad (18)$$

E. GATED RECURRENT UNIT NEURAL NETWORK

GRU neural network has been widely applied for wind speed prediction in recent years [45], [46]. As an important variant

of LSTM network, it not only inherits the ability of LSTM network to deal with nonlinear time series problems, but also simplifies the structure of LSTM network. GRU neural network improves the design of the gate. That means it combines the input gate and forgetting gate in LSTM into a single gate called the update gate and retain the original reset door. Therefore, the number of parameters is reduced and the training speed is greatly improved. The structure of GRU unit is illustrated in Fig.1. From the figure, it can be observed that z_t represents the update gate which is utilized to regulate the degree of state information from the previous moment brought into the current state while r_t represents the reset gate which is utilized to regulate the degree of state information ignored in the former moment.

The state of reset gate r_t as well as update gate z_t are defined as follows:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (19)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (20)$$

where x_t, W_r, W_z signify the input wind speed data, the weight matrixes of reset gate and update gate, respectively. Furthermore, the hidden state h_t and the candidate



FIGURE 3. Short-term wind speed time datasets with 10 min and 1 h intervals.

hidden state \tilde{h}_t can be calculated according to the following formula:

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \tag{21}$$

$$\tilde{h}_t = \tanh(W \cdot [h_{t-1} * r_t, x_t]) \tag{22}$$

where \cdot denotes point multiplication. In Equations (19), (20) and (22), two different activation functions can be determined respectively as follows:

$$\sigma(x) = \frac{1}{1 + \exp(x)} \tag{23}$$

$$\tanh(x) = \frac{1 - \exp(2x)}{1 + \exp(2x)} \tag{24}$$

III. THE HYBRID MULTI-STEP FORECASTING MODEL

The specific procedure of the proposed multi-step forecasting model is illustrated in Fig.2. and its detailed process can be explained as follows:

(1) CEEMDAN decomposition technique is employed to decompose the raw wind speed data into a sequence of $IMF_{i(i=1,2,\dots,n)}$ and a residual component named Res.

(2) SSA is utilized to capture principal components C_i' and residual components C_{ir}'' ($i = 1, 2, \dots, n$) of each IMF_i , then all the residual component C_{ir}'' and the Res are added together as a new component named C_{n+1}' . Meanwhile, among all C_i' components newly constituted, the first five components are taken as high-frequency components and the

TABLE 1. Statistical information of the four collected datasets.

Datasets	Statistic indices					
	Mean (m/s)	Max. (m/s)	Min. (m/s)	Std.	Skew. (skewness)	Kurt. (kurtosis)
SG Jan.	5.3675	14.45	0.88	2.4291	0.8482	3.672
SG Jul.	3.9591	7.92	0.35	1.7096	-0.0302	2.3256
SG Sep.	4.3686	9.63	0.55	1.5146	0.3613	3.2927
SG Dec.	7.6257	18.4	0.35	3.6208	0.2886	2.2292

remaining ones as low-frequency components according to the degree of fluctuation.

(3) PSR is utilized to construct the input and output of training set and testing set of all high-frequency and low-frequency $C_i'(i=1,2,\dots,n+1)$ acquired by CEEMDAN-SSA for the proposed model. Subsequently, KELM is employed to forecast high-frequency C_i' while GRU neural network is employed to forecast the low-frequency C_i' .

(4) The final forecasting result of the raw wind speed sequence is acquired by summing up the predicted values of each component.

IV. CASE STUDY

A. WIND SPEED DATA COLLECTION

In this section, four wind speed data datasets possessing time intervals of 10 minutes and 1 hour are collected from Sotavento Galicia (SG) wind farm. Besides, four datasets in this study are expressed as SG Jan., SG Jul., SG Sep. and SG Dec. whose date details are orderly January 01-31 in 2018 (SG Jan.), July 15-21 in 2018 (SG Jul.), September 08-14 in 2018 (SG Sep.) and December 01-31 in 2018 (SG Dec.). Furthermore, the corresponding visual presentation and statistical information are depicted in Fig.3. and table 1, respectively. Meanwhile, the statistical information includes the maximum (Max.) value, minimum (Min.) value, mean value, skewness (Skew.), kurtosis (Kurt.), and standard deviation (Std.). It can be distinctly discovered that the original wind speed data holds strong non-linearity and non-stationarity, which entails that it is hard to develop accurate prediction model. Moreover, the input and output of the training set and testing set corresponding to each predictive component are derived from PSR with different parameters, and the last 200 samples of SG Jul. and SG Sep. datasets at the 10 minutes interval are served as the testing sets while the last 150 samples of SG Jan. and SG Dec. datasets at the 1 hour interval are served as testing sets.

B. EXPERIMENTAL DESCRIPTION

1) CONTRAST MODELS AND EVALUATION INDEXES

To validate the effectiveness of the proposed hybrid model, four experiments are performed. Besides, in each experiment, several single models and compositional models are implemented as comparing models for 1-step, 3-step and 5-step prediction. The single models include SVR, KELM and GRU neural network. The remaining three comparing compositional models include CEEMDAN-GRU, CEEMDAN-SSA-PSR-KELM and CEEMDAN-SSA-PSR-GRU.

TABLE 2. Six evaluation metrics.

Metrics	Definition	Equation
$MAPE$	Absolute percentage error	$MAPE = \frac{1}{N} \sum_{i=1}^N Y_i - \hat{Y}_i / Y_i \times 100$
MAE	Mean absolute error	$MAE = \frac{1}{N} \sum_{i=1}^N Y_i - \hat{Y}_i $
$RMSE$	Root-mean-square error	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2}$
P_{MAPE}	Promoting percentages of mean absolute percentage error	$P_{MAPE} = (\frac{MAPE_a - MAPE_b}{MAPE_a}) \times 100$
P_{MAE}	Promoting percentages of mean absolute error	$P_{MAE} = (\frac{MAE_a - MAE_b}{MAE_a}) \times 100$
P_{RMSE}	Promoting percentages of root-mean-square error	$P_{RMSE} = (\frac{RMSE_a - RMSE_b}{RMSE_a}) \times 100$

In order to quantitatively assess the prediction capability of all experimental forecasting models, three accuracy indexes are employed in this study containing mean absolute percentage error (MAPE), mean absolute error (MAE) as well as root mean square error (RMSE) [47]. Besides, for purpose of comparing the improvement rates of these three indexes between the proposed prediction model and different comparison models, the corresponding improvement rates of indexes are defined and proposed. The detailed definitions and formulas of the above six evaluation indexes are presented in table 2, in which N represents the amount of predictive wind speed data, Y and \hat{Y}_i signify the actual wind speed as well as predicted values, respectively.

2) PARAMETERS SETTING OF ALL EXPERIMENTAL MODELS

It is noteworthy that the predictor GRU neural network is developed with Python 3.7.0 and Keras 2.3.1 backend with TensorFlow 2.0.0, and the rest modules are conducted with MATLAB R2018b. At the same time, Adam optimization algorithm is employed to optimize the internal parameters of GRU neural network. Besides, grid search(GS) is utilized to identify the regularization coefficient C and kernel parameters δ in SVR, KELM as well as the KELM-based comparing model, whose searching scope is $[2^{-8}, 2^8]$ and $[2^{-10}, 2^{10}]$, and both grow exponentially at 0.8 steps. Additionally, the hyper-parameters of the proposed model and comparison model involving GRU neural network are consistent to better compare the performance of corresponding prediction models, specifically including a single hidden layer, 500 times epochs of training and batch size 32. Simultaneously, Adam optimization algorithm is conducted to optimize the inherent parameters of GRU neural network. Similarly, the setting of relevant parameters of the compositional models involving CEEMDAN remains uniform, where the standard deviation of the added white noise, the number of realizations as well as the maximum number of sifting iterations are set as 0.2, 500 and 5000 in order. Besides, for all prediction models based on SSA, the embedded dimension L of the embedding phase and the grouping threshold of the grouping stage within SSA are set as 9 and 80% respectively [40].

TABLE 3. Performance evaluation of models with various prediction horizons for four datasets.

Datasets	Models	MAPE(%)			MAE(m/s)			RMSE(m/s)		
		1-step	3-step	5-step	1-step	3-step	5-step	1-step	3-step	5-step
SG Jan.	SVR	19.4309	31.4794	38.3726	0.6585	1.0844	1.2859	0.8361	1.3286	1.5484
	KELM	19.2007	31.1467	40.5418	0.6566	1.0390	1.3130	0.8322	1.2741	1.5662
	GRU	26.3808	34.3532	48.8464	0.8823	1.2013	1.6570	1.0727	1.5073	2.0588
	CEEMDAN-GRU	12.1172	17.3946	18.4165	0.4121	0.5639	0.6447	0.5413	0.7158	0.8365
	CEEMDAN-SSA-PSR-KELM	8.5536	12.7527	14.3726	0.2987	0.4025	0.4656	0.4210	0.5422	0.6131
	CEEMDAN-SSA-PSR-GRU	6.4960	13.1407	17.6526	0.2290	0.4470	0.6033	0.2908	0.5535	0.7340
	The proposed	8.4155	12.3716	14.0506	0.2958	0.4011	0.4625	0.4122	0.5334	0.6128
SG Jul.	SVR	22.8403	43.6722	59.6165	0.4722	0.7661	0.9289	0.6136	0.9634	1.1296
	KELM	23.5694	41.2231	56.0299	0.4772	0.7347	0.8998	0.6172	0.9169	1.0973
	GRU	24.4641	53.1964	66.2904	0.6421	1.0123	1.2650	0.8239	1.3717	1.6077
	CEEMDAN-GRU	8.1400	14.5530	17.8916	0.2065	0.3338	0.4117	0.2796	0.4239	0.5193
	CEEMDAN-SSA-PSR-KELM	6.6038	13.2644	17.1689	0.1425	0.2834	0.3725	0.1975	0.3668	0.5132
	CEEMDAN-SSA-PSR-GRU	6.2586	12.9513	17.5481	0.1262	0.3112	0.3654	0.1581	0.3901	0.4685
	The proposed	4.5387	11.1062	13.7212	0.1144	0.2524	0.3002	0.1685	0.3291	0.3914
SG Sep.	SVR	8.8494	14.0474	17.0545	0.5034	0.7685	0.9485	0.8955	1.0757	1.3337
	KELM	9.1200	13.7283	17.4329	0.5056	0.7397	0.9373	0.8307	1.0098	1.2243
	GRU	10.2886	17.2398	18.7871	0.5077	0.8530	0.9584	0.6815	1.1311	1.2239
	CEEMDAN-GRU	3.5156	5.4745	6.0943	0.1760	0.2763	0.2996	0.2536	0.3751	0.3882
	CEEMDAN-SSA-PSR-KELM	5.2444	8.0250	18.8467	0.3064	0.4559	1.1139	0.4781	0.6954	1.8788
	CEEMDAN-SSA-PSR-GRU	2.6986	5.2318	5.9430	0.1426	0.2631	0.2930	0.1796	0.3451	0.3613
	The proposed	2.5043	4.4602	4.9337	0.1313	0.2271	0.2541	0.1709	0.2958	0.3416
SG Dec.	SVR	27.5982	48.2372	56.4814	0.7887	1.3052	1.5879	0.9906	1.6151	1.9483
	KELM	26.6333	45.7186	57.2899	0.7718	1.2789	1.5622	0.9814	1.5943	1.9105
	GRU	33.7705	49.4284	58.9822	1.1919	1.6595	1.7822	1.4259	2.0652	2.2798
	CEEMDAN-GRU	12.1471	17.3304	20.2347	0.4649	0.6002	0.6675	0.6396	0.7650	0.8494
	CEEMDAN-SSA-PSR-KELM	7.0492	14.1614	17.2601	0.2532	0.4606	0.5671	0.3351	0.5922	0.6883
	CEEMDAN-SSA-PSR-GRU	8.1264	15.2603	18.9584	0.2606	0.4941	0.6846	0.3347	0.6315	0.8792
	The proposed	6.6150	13.8110	15.0882	0.2308	0.4458	0.4915	0.3124	0.5758	0.6171

C. EXPERIMENTAL ANALYSIS OF MULTI-STEP FORECASTING

In this part, the prediction results acquired by all the prediction models at different forecasting levels on four datasets are analyzed and discussed in detail. The indexes MAPE, MAE, and RMSE mentioned above for all the prediction models are described in Table 3. Additionally, the corresponding indexes decline rate between the proposed model as well as the comparison model are expressed in Table 4.

After the comprehensive analysis of Table 3 and Table 4, the following conclusions can be drawn as follows:

(1) Comparing the values of the single models and the combined models on the indicators MAPE, MAE and RMSE, it can be found that the compositional models combined with data preprocessing technology achieve more excellent performance on the three indicators, which indicates that the data preprocessing technology can contribute to improving the capability of the forecasting model. For instance, in the case of SG Dec. dataset, from one-step to multi-step forecasting, the MAPE values of CEEMDAN-GRU model are 12.1471%, 17.3304% and 20.2347%, respectively, while that of GRU model are 33.7705%, 49.4284% and 58.9822%, respectively, as well the same situation occurs on indicators MAE and

RMSE. Similarly, the law continues to be maintained for other three datasets.

(2) The proposed CEEMDAN-SSA-PSR-KELM-GRU model performs better multi-step forecasting capability than CEEMDAN-GRU model. For instance, in the case of SG Dec. dataset, from one-step to multi-step forecasting, the MAPE values are decreased by 45.5420%, 20.3073% and 25.4340%, respectively; the MAE values are decreased by 50.3602%, 25.7230% and 26.3663%, respectively; the RMSE values are decreased by 51.1593%, 24.7373% and 27.3538%, respectively.

(3) CEEMDAN-SSA-PSR-GRU model performs better multi-step prediction capability than that of CEEMDAN-GRU model which means that data preprocessing combined with SSA technology and PSR algorithm can further enhance the prediction performance of the model. For instance, in the case of SG Dec. dataset, from one-step to multi-step forecasting, the MAPE values of CEEMDAN-SSA-PSR-GRU model are 8.1264%, 15.2603% and 18.9584%, respectively while that of CEEMDAN-GRU model are 12.1471%, 17.3304% and 20.2347%, respectively, and the same situation occurs on indicators MAE and RMSE. Similarly, the law continues to be maintained in other three datasets.

TABLE 4. Decrement ratio of the evaluation indicators obtained by the proposed model compared with relevant contrast models.

Datasets	Models	P _{MAPE} (%)			P _{MAE} (%)			P _{RMSE} (%)		
		1-step	3-step	5-step	1-step	3-step	5-step	1-step	3-step	5-step
SG Jan.	SVR	56.6903	60.6996	63.3839	55.0834	63.0065	64.0332	50.7004	59.8544	60.4211
	KELM	56.1710	60.2797	65.3430	54.9553	61.3900	64.7751	50.4674	58.1364	60.8729
	GRU	68.1000	63.9872	71.2352	66.4778	66.6062	72.0882	61.5716	64.6137	70.2345
	CEEMDAN-GRU	30.5492	28.8770	23.7066	28.2241	28.8607	28.2588	23.8460	25.4826	26.7374
	CEEMDAN-SSA-PSR-KELM	1.6153	2.9888	2.2403	0.9668	0.3332	0.6651	2.0835	1.6248	0.0499
	CEEMDAN-SSA-PSR-GRU	-29.5479	5.8535	20.4053	-29.1724	10.2582	23.3317	-41.7461	3.6447	16.5039
SG Jul.	SVR	80.1285	74.5691	76.9843	75.7617	67.0531	67.6817	72.5356	65.8348	65.3486
	KELM	80.7432	73.0582	75.5110	76.0152	65.6469	66.6396	72.6939	64.1033	64.3279
	GRU	81.4474	79.1222	79.3014	82.1766	75.0651	76.2702	79.5444	76.0046	75.6528
	CEEMDAN-GRU	44.2418	23.6841	23.3094	44.5712	24.3742	27.0764	39.7165	22.3484	24.6257
	CEEMDAN-SSA-PSR-KELM	31.2716	16.2703	20.0813	19.6677	10.9476	19.4116	14.6656	10.2723	23.7337
	CEEMDAN-SSA-PSR-GRU	27.4807	14.2464	21.8082	9.2892	18.8980	17.8438	-6.6118	15.6342	16.4548
SG Sep.	SVR	71.7013	68.2492	71.0712	73.9094	70.4459	73.2127	80.9182	72.5020	74.3887
	KELM	72.5411	67.5113	71.6991	74.0212	69.2947	72.8943	79.4286	70.7052	72.0999
	GRU	75.6598	74.1287	73.7391	74.1320	73.3753	73.4912	74.9266	73.8474	72.0912
	CEEMDAN-GRU	28.7680	18.5284	19.0452	25.3539	17.8000	15.1964	32.6137	21.1303	12.0083
	CEEMDAN-SSA-PSR-KELM	52.2491	44.4220	73.8221	57.1267	50.1829	77.1919	64.2557	57.4607	81.8193
	CEEMDAN-SSA-PSR-GRU	7.2004	14.7494	16.9832	7.8857	13.6669	13.2936	4.8277	14.2839	5.4601
SG Dec.	SVR	76.0309	71.3685	73.2863	70.7362	65.8471	69.0468	68.4642	64.3496	68.3266
	KELM	75.1625	69.7912	73.6633	70.0975	65.1449	68.5374	68.1674	63.8856	67.6992
	GRU	80.4118	72.0585	74.4190	80.6361	73.1382	72.4212	78.0907	72.1192	72.9323
	CEEMDAN-GRU	45.5420	20.3073	25.4340	50.3602	25.7230	26.3663	51.1593	24.7373	27.3538
	CEEMDAN-SSA-PSR-KELM	6.1594	2.4743	12.5833	8.8595	3.2159	13.3328	6.7609	2.7762	10.3464
	CEEMDAN-SSA-PSR-GRU	18.5982	9.4970	20.4141	11.4288	9.7781	28.2018	6.6631	8.8272	29.8123

(4) The proposed CEEMDAN-SSA-PSR-KELM-GRU model performs better multi-step forecasting capability than the CEEMDAN-SSA-PSR-KELM model, which demonstrates that the proposed hybrid model based on two predictors can achieve higher multi-step prediction performance. For instance, in the case of SG Dec. dataset, from one-step to multi-step forecasting, the MAPE values are decreased by 6.1594%, 2.4743% and 12.5833%, respectively; the MAE values are decreased by 8.8595%, 3.2159% and 13.3328%, respectively; the indexes RMSE are decreased by 6.7609%, 2.7762% and 10.3464%, respectively.

(5) The proposed CEEMDAN-SSA-PSR-KELM-GRU model performs better multi-step forecasting capability than the CEEMDAN-SSA-PSR-GRU model. For instance, in the case of SG Dec. dataset, from one-step to multi-step, the MAPE values are decreased by 18.5982%, 9.4970% and 20.4141%, respectively; the MAE values are decreased by 11.4288%, 9.7781% and 28.2018%, respectively; the RMSE values are decreased by 6.6631%, 8.8272% and 29.8123%, respectively.

(6) The CEEMDAN-SSA-PSR-GRU model achieves high prediction accuracy in the prediction of 1-step of SG Jan. dataset and the prediction accuracy of CEEMDAN-SSA-PSR-KELM is similar to the proposed model. This indicates that compared with other models, the proposed model removes some detail components and extracts the principal components in the wind speed sequence. However, in all experimental models, the proposed multi-step forecasting model obtains the optimal results.

D. VISUAL ANALYSIS OF FORECASTING RESULTS

In this part, the visual presentation of prediction results of the experiment on SG Dec. dataset is taken as an example for discussing separately. The forecasted curves of all experimental models and the bar diagram of actual wind speed values as well as scatter diagrams describing the degree of fit for each model are illustrated in Fig.4. (a)-(c). From the Fig.4, it can be clearly observed that with the increasement of prediction level, the fitting curves of single models will diverge from the actual wind speed values while the forecasting curves of compositional models based on decomposition are relatively stable. Therefore, the conclusions (4) and (5) can be drawn more evidently from Fig.4. and the fitting curves of different prediction levels derived from the proposed model are closer to the actual wind speed values region. Furthermore, according to the scatter diagrams of different forecasting levels, it is clearly that the predicted values of actual wind speed for the proposed model are distributed most uniformly on the regression line as well as the index R of the proposed model is the optimal in all experimental models, i.e., 0.98686, 0.95689 and 0.94992, respectively. Namely, compared with other comparing models, the prediction accuracy of the proposed model would not decrease markedly with the increasement of the forecasting level. Therefore, the predictive capability of the proposed model in multi-step short-term wind speed forecasting has been well verified

Moreover, the evaluation indexes illustrated in Table 3 are visually demonstrated in Fig.5 with radar charts and each indicator implemented at different forecasting levels is

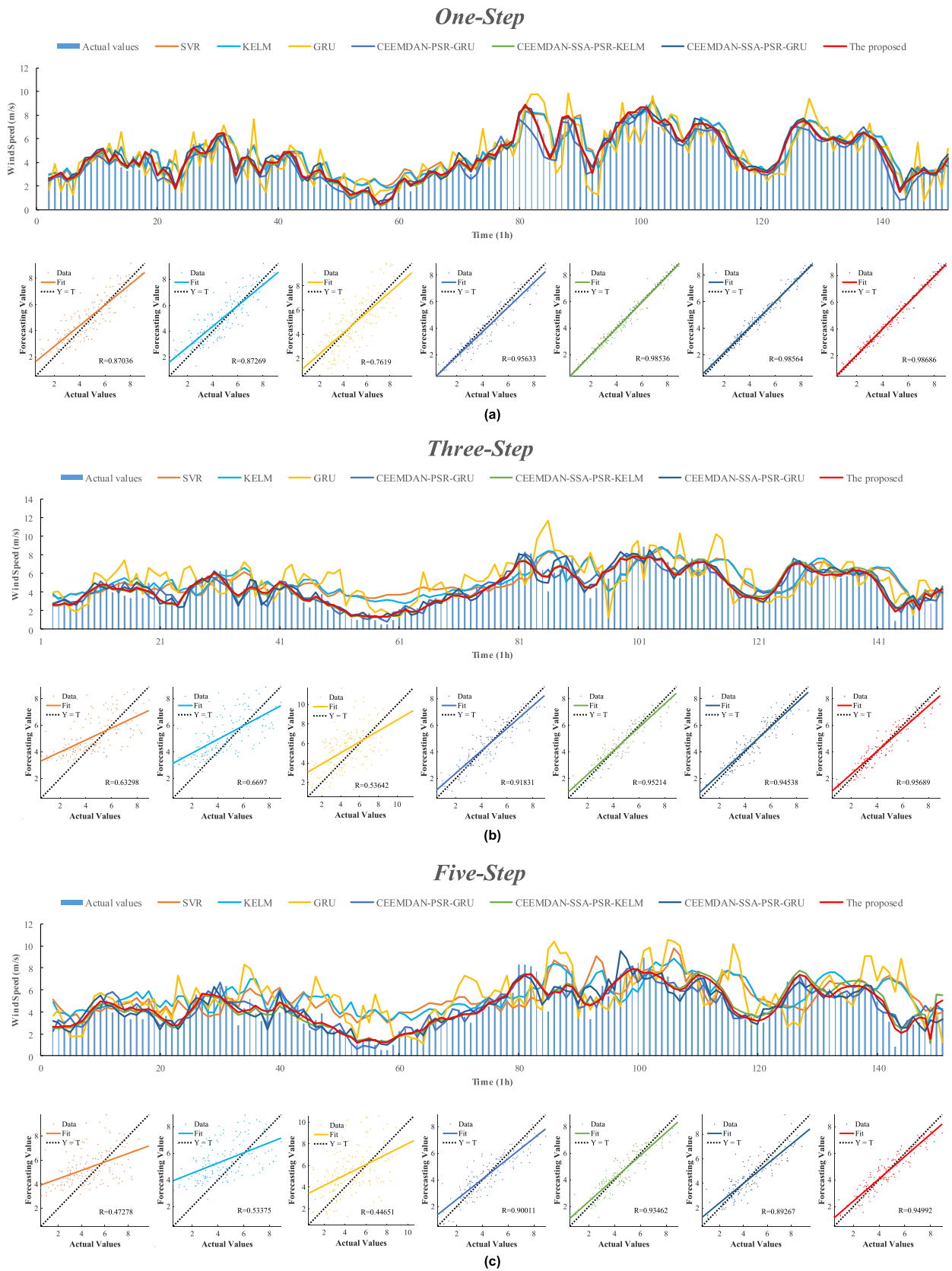


FIGURE 4. Multi-step forecasting results of all models for dataset of SG Dec.: (a) one-step, (b) three-step and (c) five-step.

explained successively in the same subgraph. Therefore, it can be clearly observed that the variation trends of indexes

with the increasement of prediction levels. To be more specific, it can be concluded from the forecasting results of all



FIGURE 5. Radar charts for visualized observation of the indexes and the variation tendency in different prediction levels obtained by all models: (a)SG Jan. (b)SG Jul. (c)SG Sep. (d)SG Dec.

models for dataset SG Dec. as shown in figure 5(a) that the three indexes representing the prediction performance are all minimum, as well the variation range of each index is small.

V. CONCLUSION

In order to establish an accurate multi-step short-term wind speed forecasting model, the hybrid model including multi-stage principal component extraction, GRU neural network and KELM is proposed in this paper. Among which, multi-stage principal component extraction combines CEEMDAN, SSA and PSR. Firstly, CEEMDAN is employed to decompose the raw wind speed data into a sequence of IMFs and a residual component. Then, SSA is adopted to further capture dominant and residual components of all IMFs and all residual components obtained by CEEMDAN decomposition and SSA processing are added to form a

new predicted component. Subsequently, PSR is utilized to construct each component obtained by CEEMDAN-SSA into the input and output of training set and testing set for the prediction model. Following, GRU neural network is employed to forecast the low-frequency components while KELM is employed to forecast the high-frequency components. Finally, the predicted values of all components are accumulated to acquire the ultimate prediction result of the raw wind speed data with different prediction levels. To verify the multi-step forecasting capability of the proposed model, some comparing models including SVR, KELM, GRU neural network, CEEMDAN-PSR-GRU, CEEMDAN-SSA-PSR-KELM as well as CEEMDAN-SSA-PSR-GRU are employed for comparison. The experimental results of four datasets demonstrate that the proposed hybrid model can achieve better multi-step wind speed prediction effect.

However, there still exists some limitations in this study, such as the quality of the raw wind speed data as well as the selection of relevant parameters of the experimental models according to the research results of other scholars. In future research work, we will develop effective methods to further improve the quality of raw wind speed data, and utilize corresponding optimization methods to optimize model parameters.

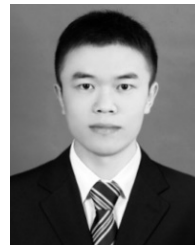
REFERENCES

- [1] X. Mi, H. Liu, and Y. Li, "Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine," *Energy Convers. Manage.*, vol. 180, pp. 196–205, Jan. 2019, doi: [10.1016/j.enconman.2018.11.006](https://doi.org/10.1016/j.enconman.2018.11.006).
- [2] J. Wang, W. Yang, P. Du, and T. Niu, "A novel hybrid forecasting system of wind speed based on a newly developed multi-objective sine cosine algorithm," *Energy Convers. Manage.*, vol. 163, pp. 134–150, May 2018, doi: [10.1016/j.enconman.2018.02.012](https://doi.org/10.1016/j.enconman.2018.02.012).
- [3] W. Fu, K. Wang, C. Li, and J. Tan, "Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM," *Energy Convers. Manage.*, vol. 187, pp. 356–377, May 2019, doi: [10.1016/j.enconman.2019.02.086](https://doi.org/10.1016/j.enconman.2019.02.086).
- [4] 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, doi: [10.1109/ACCESS.2020.3011060](https://doi.org/10.1109/ACCESS.2020.3011060).
- [5] M. Lei, L. Shiyuan, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renew. Sustain. Energy Rev.*, vol. 13, no. 4, pp. 915–920, May 2009, doi: [10.1016/j.rser.2008.02.002](https://doi.org/10.1016/j.rser.2008.02.002).
- [6] J. Zhao, Z.-H. Guo, Z.-Y. Su, Z.-Y. Zhao, X. Xiao, and F. Liu, "An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed," *Appl. Energy*, vol. 162, pp. 808–826, Jan. 2016, doi: [10.1016/j.apenergy.2015.10.145](https://doi.org/10.1016/j.apenergy.2015.10.145).
- [7] J. Zhang, C. Draxl, T. Hopson, L. D. Monache, E. Vanvyve, and B.-M. Hodge, "Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods," *Appl. Energy*, vol. 156, pp. 528–541, Oct. 2015, doi: [10.1016/j.apenergy.2015.07.059](https://doi.org/10.1016/j.apenergy.2015.07.059).
- [8] W. Zhao, Y.-M. Wei, and Z. Su, "One day ahead wind speed forecasting: A resampling-based approach," *Appl. Energy*, vol. 178, pp. 886–901, Sep. 2016, doi: [10.1016/j.apenergy.2016.06.098](https://doi.org/10.1016/j.apenergy.2016.06.098).
- [9] G. H. Riahy and M. Abedi, "Short term wind speed forecasting for wind turbine applications using linear prediction method," *Renew. Energy*, vol. 33, no. 1, pp. 35–41, Jan. 2008, doi: [10.1016/j.renene.2007.01.014](https://doi.org/10.1016/j.renene.2007.01.014).
- [10] Y. Noorollahi, M. A. Jokar, and A. Kalhor, "Using artificial neural networks for temporal and spatial wind speed forecasting in iran," *Energy Convers. Manage.*, vol. 115, pp. 17–25, May 2016, doi: [10.1016/j.enconman.2016.02.041](https://doi.org/10.1016/j.enconman.2016.02.041).
- [11] P. Poggi, M. Muselli, G. Notton, C. Cristofari, and A. Louche, "Forecasting and simulating wind speed in Corsica by using an autoregressive model," *Energy Convers. Manage.*, vol. 44, no. 20, pp. 3177–3196, 2003, doi: [10.1016/S0196-8904\(03\)00108-0](https://doi.org/10.1016/S0196-8904(03)00108-0).
- [12] E. Erdem and J. Shi, "ARMA based approaches for forecasting the tuple of wind speed and direction," *Appl. Energy*, vol. 88, no. 4, pp. 1405–1414, Apr. 2011, doi: [10.1016/j.apenergy.2010.10.031](https://doi.org/10.1016/j.apenergy.2010.10.031).
- [13] H. Liu, H. Tian, and Y. Li, "An EMD-recursive ARIMA method to predict wind speed for railway strong wind warning system," *J. Wind Eng. Ind. Aerodyn.*, vol. 141, pp. 27–38, Jun. 2015, doi: [10.1016/j.jweia.2015.02.004](https://doi.org/10.1016/j.jweia.2015.02.004).
- [14] R. G. Kavasseri and K. Seetharaman, "Day-ahead wind speed forecasting using f-ARIMA models," *Renew. Energy*, vol. 34, no. 5, pp. 1388–1393, May 2009, doi: [10.1016/j.renene.2008.09.006](https://doi.org/10.1016/j.renene.2008.09.006).
- [15] O. Ait Maatallah, A. Achuthan, K. Janoyan, and P. Marzocca, "Recursive wind speed forecasting based on hammerstein autoregressive model," *Appl. Energy*, vol. 145, pp. 191–197, May 2015, doi: [10.1016/j.apenergy.2015.02.032](https://doi.org/10.1016/j.apenergy.2015.02.032).
- [16] Q. Han, F. Meng, T. Hu, and F. Chu, "Non-parametric hybrid models for wind speed forecasting," *Energy Convers. Manage.*, vol. 148, pp. 554–568, Sep. 2017, doi: [10.1016/j.enconman.2017.06.021](https://doi.org/10.1016/j.enconman.2017.06.021).
- [17] L. Xiao, J. Wang, X. Yang, and L. Xiao, "A hybrid model based on data preprocessing for electrical power forecasting," *Electr. Power Energy Syst.*, vol. 64, pp. 311–327, Jan. 2015, doi: [10.1016/j.ijepes.2014.07.029](https://doi.org/10.1016/j.ijepes.2014.07.029).
- [18] Z.-H. Guo, J. Wu, H.-Y. Lu, and J.-Z. Wang, "A case study on a hybrid wind speed forecasting method using BP neural network," *Knowl.-Based Syst.*, vol. 24, no. 7, pp. 1048–1056, Oct. 2011, doi: [10.1016/j.knsys.2011.04.019](https://doi.org/10.1016/j.knsys.2011.04.019).
- [19] W. He, Z. Wang, and H. Jiang, "Model optimizing and feature selecting for support vector regression in time series forecasting," *Neurocomputing*, vol. 72, nos. 1–3, pp. 600–611, Dec. 2008, doi: [10.1016/j.neucom.2007.11.010](https://doi.org/10.1016/j.neucom.2007.11.010).
- [20] K. Shao, W. Fu, J. Tan, and K. Wang, "Coordinated approach fusing time-shift multiscale dispersion entropy and vibrational Harris hawks optimization-based SVM for fault diagnosis of rolling bearing," *Measurement*, Oct. 2020, Art. no. 108580, doi: [10.1016/j.measurement.2020.108580](https://doi.org/10.1016/j.measurement.2020.108580).
- [21] Y. Zhao, L. Ye, Z. Li, X. Song, Y. Lang, and J. Su, "A novel bidirectional mechanism based on time series model for wind power forecasting," *Appl. Energy*, vol. 177, pp. 793–803, Sep. 2016, doi: [10.1016/j.apenergy.2016.03.096](https://doi.org/10.1016/j.apenergy.2016.03.096).
- [22] S. Zhu, X. Luo, Z. Xu, and L. Ye, "Seasonal streamflow forecasts using mixture-kernel GPR and advanced methods of input variable selection," *Hydrol. Res.*, vol. 50, pp. 1–15, Feb. 2018, doi: [10.2166/nh.2018.023](https://doi.org/10.2166/nh.2018.023).
- [23] S. Zhu, X. Yuan, Z. Xu, X. Luo, and H. Zhang, "Gaussian mixture model coupled recurrent neural networks for wind speed interval forecast," *Energy Convers. Manage.*, vol. 198, Oct. 2019, Art. no. 111772, doi: [10.1016/j.enconman.2019.06.083](https://doi.org/10.1016/j.enconman.2019.06.083).
- [24] J. Zhou, J. Shi, and G. Li, "Fine tuning support vector machines for short-term wind speed forecasting," *Energy Convers. Manage.*, vol. 52, no. 4, pp. 1990–1998, Apr. 2011, doi: [10.1016/j.enconman.2010.11.007](https://doi.org/10.1016/j.enconman.2010.11.007).
- [25] C. Yu, Y. Li, and M. Zhang, "Comparative study on three new hybrid models using elman neural network and empirical mode decomposition based technologies improved by singular spectrum analysis for hour-ahead wind speed forecasting," *Energy Convers. Manage.*, vol. 147, pp. 75–85, Sep. 2017, doi: [10.1016/j.enconman.2017.05.008](https://doi.org/10.1016/j.enconman.2017.05.008).
- [26] Z. A. Bashir and M. E. El-Hawary, "Applying wavelets to short-term load forecasting using PSO-based neural networks," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 20–27, Feb. 2009, doi: [10.1109/TPWRS.2008.2008606](https://doi.org/10.1109/TPWRS.2008.2008606).
- [27] G. Xu, C. Xiu, and Z. Wan, "Hysteretic chaotic operator network and its application in wind speed series prediction," *Neurocomputing*, vol. 165, pp. 384–388, Oct. 2015, doi: [10.1016/j.neucom.2015.03.027](https://doi.org/10.1016/j.neucom.2015.03.027).
- [28] G. Chen, X. Yi, Z. Zhang, and H. Wang, "Applications of multi-objective dimension-based firefly algorithm to optimize the power losses, emission, and cost in power systems," *Appl. Soft Comput.*, vol. 68, pp. 322–342, Jul. 2018, doi: [10.1016/j.asoc.2018.04.006](https://doi.org/10.1016/j.asoc.2018.04.006).
- [29] K. Wang, W. Fu, T. Chen, B. Zhang, D. Xiong, and P. Fang, "A compound framework for wind speed forecasting based on comprehensive feature selection, quantile regression incorporated into convolutional simplified long short-term memory network and residual error correction," *Energy Convers. Manage.*, vol. 222, Oct. 2020, Art. no. 113234, doi: [10.1016/j.enconman.2020.113234](https://doi.org/10.1016/j.enconman.2020.113234).
- [30] N. Chen, Z. Qian, and X. Meng, "Multistep wind speed forecasting based on wavelet and Gaussian processes," *Math. Problems Eng.*, vol. 2013, pp. 1–8, Aug. 2013, doi: [10.1155/2013/461983](https://doi.org/10.1155/2013/461983).
- [31] H. Liu, C. Chen, H.-Q. Tian, and Y.-F. Li, "A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks," *Renew. Energy*, vol. 48, pp. 545–556, Dec. 2012, doi: [10.1016/j.renene.2012.06.012](https://doi.org/10.1016/j.renene.2012.06.012).
- [32] H. Liu, H.-Q. Tian, D.-F. Pan, and Y.-F. Li, "Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks," *Appl. Energy*, vol. 107, pp. 191–208, Jul. 2013, doi: [10.1016/j.apenergy.2013.02.002](https://doi.org/10.1016/j.apenergy.2013.02.002).
- [33] Y. Wang, S. Wang, and N. Zhang, "A novel wind speed forecasting method based on ensemble empirical mode decomposition and GA-BP neural network," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, vol. 94, Jan. 2013, pp. 1–5, doi: [10.1109/PESMG.2013.6672195](https://doi.org/10.1109/PESMG.2013.6672195).
- [34] G. Poitras and G. Cormier, "Wind speed prediction for a target station using neural networks and particle swarm optimization," *Wind Eng.*, vol. 35, no. 3, pp. 369–380, Jun. 2011, doi: [10.1260/0309-524X.35.3.369](https://doi.org/10.1260/0309-524X.35.3.369).
- [35] H. Liu, H.-Q. Tian, C. Chen, and Y.-F. Li, "An experimental investigation of two wavelet-MLP hybrid frameworks for wind speed prediction using GA and PSO optimization," *Int. J. Electr. Power Energy Syst.*, vol. 52, pp. 161–173, Nov. 2013, doi: [10.1016/j.ijepes.2013.03.034](https://doi.org/10.1016/j.ijepes.2013.03.034).

- [36] W. Fu and Q. Lu, "Multiobjective optimal control of FOPID controller for hydraulic turbine governing systems based on reinforced multiobjective harris hawks optimization coupling with hybrid strategies," *Complexity*, vol. 2020, Jul. 2020, Art. no. 9274980.
- [37] C. Zhang, H. Wei, J. Zhao, T. Liu, T. Zhu, and K. Zhang, "Short-term wind speed forecasting using empirical mode decomposition and feature selection," *Renew. Energy*, vol. 96, pp. 727–737, Oct. 2016, doi: 10.1016/j.renene.2016.05.023.
- [38] C. Yu, Y. Li, and M. Zhang, "An improved wavelet transform using singular spectrum analysis for wind speed forecasting based on elman neural network," *Energy Convers. Manage.*, vol. 148, pp. 895–904, Sep. 2017, doi: 10.1016/j.enconman.2017.05.063.
- [39] H. Liu, X. Mi, and Y. Li, "Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM," *Energy Convers. Manage.*, vol. 159, pp. 54–64, Mar. 2018, doi: 10.1016/j.enconman.2018.01.010.
- [40] W. Fu, K. Wang, J. Tan, and K. Zhang, "A composite framework coupling multiple feature selection, compound prediction models and novel hybrid swarm optimizer-based synchronization optimization strategy for multi-step ahead short-term wind speed forecasting," *Energy Convers. Manage.*, vol. 205, Feb. 2020, Art. no. 112461, doi: 10.1016/j.enconman.2019.112461.
- [41] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 4144–4147, doi: 10.1109/ICASSP.2011.5947265.
- [42] W. Fu, K. Wang, J. Tan, and K. Shao, "Vibration tendency prediction approach for hydropower generator fused with multiscale dominant ingredient chaotic analysis, adaptive mutation grey wolf optimizer, and KELM," *Complexity*, vol. 2020, Jan. 2020, Art. no. 4516132.
- [43] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, "Geometry from a time series," *Phys. Rev. Lett.*, vol. 45, no. 9, pp. 712–716, Sep. 1980, doi: 10.1103/PhysRevLett.45.712.
- [44] 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, doi: 10.1016/j.neucom.2005.12.126.
- [45] R. Wang, C. Li, W. Fu, and G. Tang, "Deep learning method based on gated recurrent unit and variational mode decomposition for short-term wind power interval prediction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 10, pp. 3814–3827, Oct. 2020.
- [46] C. Li, G. Tang, X. Xue, A. Saeed, and X. Hu, "Short-term wind speed interval prediction based on ensemble GRU model," *IEEE Trans. Sustain. Energy*, vol. 11, no. 3, pp. 1370–1380, Jul. 2020.
- [47] W. Fu, K. Zhang, K. Wang, B. Wen, P. Fang, and F. Zou, "A hybrid approach for multi-step wind speed forecasting based on two-layer decomposition, improved hybrid DE-HHO optimization and KELM," *Renew. Energy*, vol. 164, pp. 211–229, Feb. 2021, doi: 10.1016/j.renene.2020.09.078.



FENG ZOU is currently pursuing the master's degree in electrical engineering with China Three Gorges University. His research interests include machine learning, signal processing, and wind speed prediction.



WENLONG FU (Member, IEEE) received the B.S. and Ph.D. degrees in hydraulic and hydropower engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2011 and 2016, respectively. He is currently an Associate Professor with the College of Electrical Engineering and New Energy, China Three Gorges University. His research interests include machine learning, signal processing, and wind speed prediction.



PING FANG is currently pursuing the master's degree in electrical engineering with China Three Gorges University. Her research interests include machine learning, signal processing, and wind speed prediction.



DONGZHEN XIONG is currently pursuing the master's degree in electrical engineering with China Three Gorges University. His research interests include machine learning, signal processing, and wind speed prediction.



RENMING WANG received the Ph.D. degree in control theory and control engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2003. He is currently a Professor with the College of Electrical Engineering and New Energy, China Three Gorges University. His research interests include machine learning and power system modeling.

...