Energy Conversion and Management 143 (2017) 410–430

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm

Liye Xiao [a], Feng Qian [b], Wei Shao [a,*]

[a] School of Physical Electronics, University of Electronic Science and Technology of China, Chengdu, China
[b] Department of Electronics Engineering and Computer Science, Peking University, Beijing, China

ARTICLE INFO

Article history:
Received 17 February 2017
Received in revised form 27 March 2017
Accepted 3 April 2017

Keywords:
Hybrid forecasting architecture
Improved bat algorithm
Singular spectrum analysis, wind speed forecasting

ABSTRACT

As one of the most promising sustainable energy sources, wind energy plays an important role in energy development because of its cleanliness without causing pollution. Generally, wind speed forecasting, which has an essential influence on wind power systems, is regarded as a challenging task. Analyses based on single-step wind speed forecasting have been widely used, but their results are insufficient in ensuring the reliability and controllability of wind power systems. In this paper, a new forecasting architecture based on decomposing algorithms and modified neural networks is successfully developed for multi-step wind speed forecasting. Four different hybrid models are contained in this architecture, and to further improve the forecasting performance, a modified bat algorithm (BA) with the conjugate gradient (CG) method is developed to optimize the initial weights between layers and thresholds of the hidden layer of neural networks. To investigate the forecasting abilities of the four models, the wind speed data collected from four different wind power stations in Penglai, China, were used as a case study. The numerical experiments showed that the hybrid model including the singular spectrum analysis and general regression neural network with CG-BA (SSA-CG-BA-GRNN) achieved the most accurate forecasting results in one-step to three-step wind speed forecasting.

1. Introduction

As one of the most promising potential renewable energy sources [1], wind energy has attracted the focus of many researchers and scientists [2], and nearly every government across the world has introduced positive policies to support wind energy development [3,4]. In 2015, the global total capacity of wind farms is approximately 432,419 MW, with the 22% growth rate, as shown in Fig. 1 [5]. With the increased proportion of wind energy in whole energy networks, accurate wind speed forecasting results are becoming increasingly crucial for managers to schedule the daily power distribution and decrease the reserve capacity. To protect wind power from the breakdown and make sure the success of wind power conversation, accurate forecasting results of wind speed are also required [6]. However, due to the non-stationary and nonlinear fluctuations, wind speed is regarded as one of the hardest weather parameters to predict [7,8].

In recent decades, many methods have been presented for wind speed forecasting, and these methods can be divided into four categories [9]: (a) physical models; (b) statistical models; (c) spatial correlation models; and (d) artificial intelligence models. Physical models which are based on physical parameters, such as topography, temperature and pressure, are usually applied in long term wind speed forecasting [10–12]. Statistical models are built based on the mature statistical equations to get the potential change rule from history data sampling [13–17]. Spatial correlation models mainly consider the spatial relationship of wind speed at different sites. In some situations, it can obtain higher precision [18,19]. With the rapid development of artificial techniques, some artificial intelligence forecasting methods, including artificial neural networks (ANNs) [20–25], fuzzy logic methods [18,26] and support

Abbreviations: ANN, artificial neural network; ARIMA, autoregressive integrated moving average; BA, bat algorithm; CSA, cuckoo search algorithm; CG, conjugate gradient; EA, evolutionary algorithm; EEMD, ensemble empirical mode decomposition; EMD, empirical mode decomposition; FEEMD, fast ensemble empirical mode decomposition; FVD, forecasting validity degree; GA, genetic algorithm; GRNN, general regression neural network; MAE, mean absolute error; MAPE, mean absolute percentage error; MSE, mean square error; PSO, particle swarm optimization; RBFNN, radical basis function neural network; SDA, steepest descent algorithm; SSA, singular spectrum analysis; SVM, support vector machine; WD, wavelet decomposition; WPD, wavelet packet decomposition.

* Corresponding author at: School of Physical Electronics, University of Electronic Science and Technology of China, Chengdu 610054, China.

E-mail address: weishao@uestc.edu.cn (W. Shao).

http://dx.doi.org/10.1016/j.enconman.2017.04.012
0196-8904/

Nomenclature

α	a random vector, with a value between 0 and 1	M	total number of CG iterations
β	a random vector, with a value between 0 and 1	N	number of generations P
$\boldsymbol{d}_i^k$	the search direction of $\boldsymbol{x}_i$ at iteration k	O	output of RBFNN
d_i	Euclidian distance	r	pulse rate of a bat
ε	a random vector, with a value between 0 and 1	R	spread parameter
$\phi(d_i)$	outputs from the hidden layer of RBFNN	$r_n(t)$	nth residue
F_i	the fitness function of $\boldsymbol{x}_i$	σ	spread parameter
$-\nabla f(\boldsymbol{x}^{iter})$	gradient of $\boldsymbol{x}^{iter}$	σ_j^i	marginal standard deviation
$\boldsymbol{g}_i^k$	gradient of $\boldsymbol{x}_i$ at iteration k	S_s	simple summation
h_j^i	correlation coefficient	S_w	weighted summation
I	input vector of RBFNN	t	current iteration number
$\mathbf{IMF}_j\,(t)$	intrinsic mode function	v_i^t	the velocity of $\boldsymbol{x}_i$ at iteration t
$iter$	current iteration number	w	interconnection weight
$Iter_{\mathbf{max}}$	maximum number of iterations	w_i	weight of the hidden layer of RBFNN
K	shape matrix	X	input vector of GRNN
k_j^i	(i, j)th element of the shape matrix K	$\boldsymbol{x}_b$	the value of $\boldsymbol{x}$ with the best fitness value in the population
L	the loudness of a bat	$\boldsymbol{x}_t$	training data
m_i	center vector	$\boldsymbol{x}_i^t$	position of $\boldsymbol{x}_i$ at iteration t
m_i^j	jth element of center vector	$\boldsymbol{x}^{iter}$	positions of bats
λ_j^{iter}	step length	Y	output vector of GRNN
λ^k	step length of $\boldsymbol{x}_i$ at iteration k		

vector machines (SVMs) [27], have been developed for wind speed forecasting.

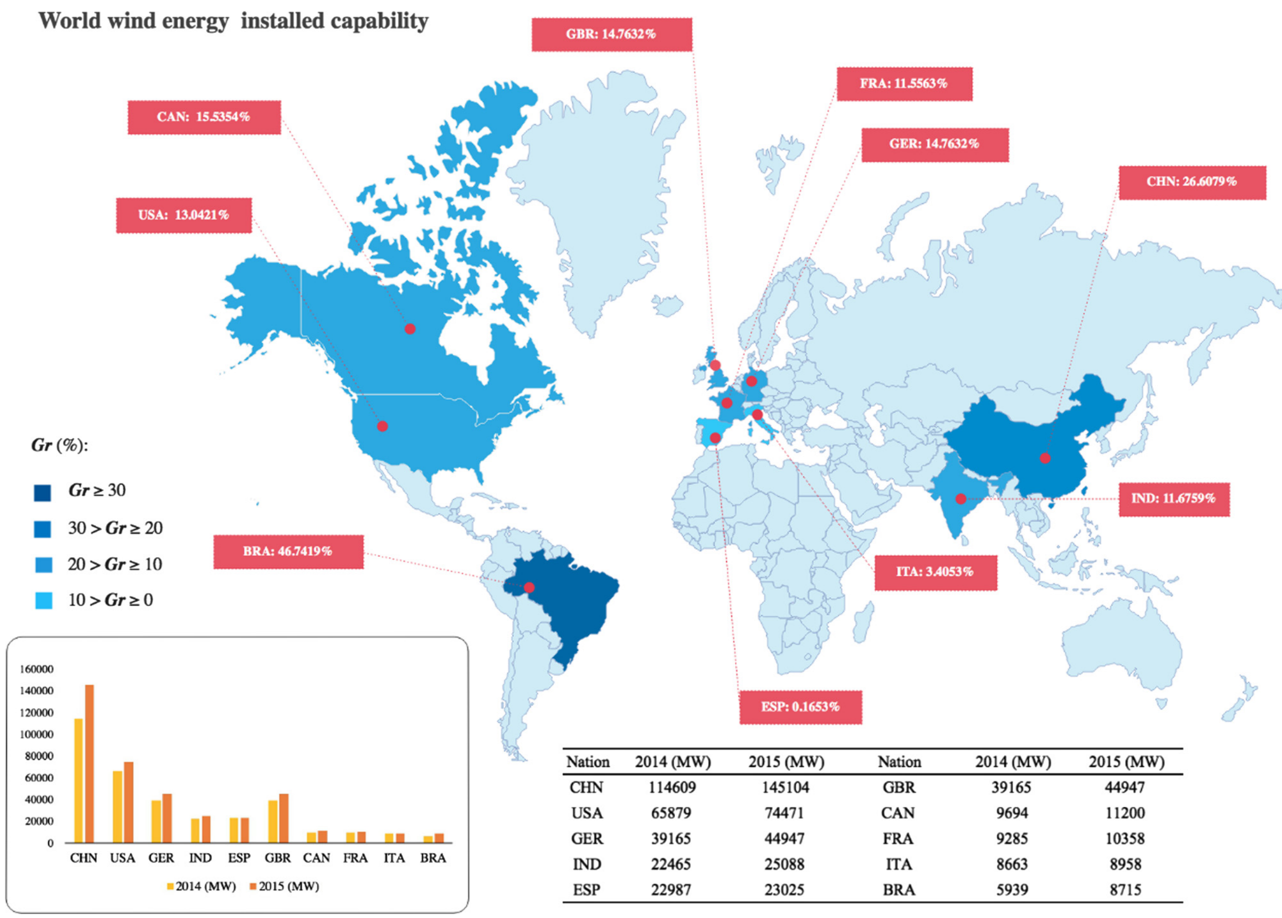
Meanwhile, to decrease the negative influences that are intrinsic to individual models, many hybrid wind speed forecasting models have been proposed [28–36].

To achieve higher forecasting accuracy, some data-processing algorithms, such as wavelet decomposition (WD) [28], wavelet packet decomposition (WPD) [29], empirical mode decomposition (EMD) [30], the ensemble empirical mode decomposition (EEMD) algorithm [31] and the fast ensemble empirical mode decomposition (FEEMD) algorithm [32], have been employed in ANNs to build hybrid models. The data decomposition, which could reduce the non-stationary feature of the original data, promotes the forecasting performance indirectly.

Moreover, intelligent optimization algorithms including the genetic algorithm (GA) [33], particle swarm optimization (PSO) [34], the evolutionary algorithm (EA) [35], and the cuckoo search algorithm (CSA) [36], are utilized to determine the initial weights and thresholds of ANNs. In 2010, Yang proposed the bat algorithm (BA) [37], which is inspired by the echolocation characteristics of bats with varying pulse rates of emission and loudness. It has been applied to a wide range of optimization applications [38], including image processing [39], classifications [40], scheduling [41], the electricity market [42], energy systems [43] and various other problems. Experiments have shown its promising efficiency for global optimization.

Analyses based on single-step wind speed forecasting have been widely used, while their results are insufficient in ensuring the reliability and controllability of wind power systems. Thus, it is required to build a model to achieve accurate results for multi-step wind speed forecasting. Among various ANN models, the radical basis function neural network (RBFNN) and general regression neural network (GRNN) are good choices to achieve high convergence rates and accurate results. In this paper, a hybrid architecture, which contains four hybrid models, with two decomposing algorithms (i.e., FEEMD and singular spectrum analysis (SSA)) which are used to realize the non-stationary wind speed decomposition, and the modified RBFNN and GRNN is proposed for wind speed forecasting. In the modified RBFNN and GRNN, an improved BA, which is on the basis of conjugate gradient (CG) method to improve convergence performance over time and prevent individual bats from entrapment in local optima, is introduced to optimize the initial weights and thresholds of RBFNN and GRNN. The aim of this study is to investigate and enhance the forecasting performance of hybrid model based on signal processing algorithms, intelligent optimization algorithm and artificial neural networks for multi-step accurate wind speed forecasting. To investigate the forecasting abilities of the four models, the wind speed data collected from four different wind power stations in Penglai, China, were used as a case study. The main contributions in this paper are demonstrated as follows.

(1) ***The forecasting focus of the forecasting architecture is not only on the single-step forecasting but also on the multi-step forecasting***. Although the wind speed single-step predictions have been studied widely, to protect the wind power, wind speed single-step forecasting results alone are insufficient, and wind speed multi-step forecasting results are definitely expected, thus the forecasting architecture is aim to enhance the forecasting accuracy of multi-step wind speed forecasting.
(2) ***To globally investigate the forecasting performance of different combination of decomposing algorithms and neural networks, a forecasting architecture contains four hybrid models is proposed***. In the architecture, four different hybrid forecasting models based on the two most popular decomposing algorithms, an improved optimization algorithm and two neural networks, are investigated and compared (the performance of multi-step forecasting is given special attention in the investigation) with four different sites data for one-step to three-step forecasting to obtain the best one.
(3) ***The speed of local convergence and the accuracy of finding the optimal solution of BA are enhanced***. To improve both the exploration and exploitation capacities and avoid the weakness of the local optima searching ability, the improved BA based on CG is proposed, and to evaluate the improved algorithm, four testing functions are used.



Nation	2014 (MW)	2015 (MW)	Nation	2014 (MW)	2015 (MW)
CHN	114609	145104	GBR	39165	44947
USA	65879	74471	CAN	9694	11200
GER	39165	44947	FRA	9285	10358
IND	22465	25088	ITA	8663	8958
ESP	22987	23025	BRA	5939	8715

Fig. 1. Top 10 countries of wind power newly increased installed capability in 2015.

(4) ***The forecasting accuracy and stability of RBFNN and GRNN are enhanced***. The improved BA, CG-BA, is employed to select the initial weights and thresholds for the RBFNN and GRNN. According to the experiment results, the forecasting performance of RBFNN and GRNN are directly enhanced with CG-BA.

(5) ***To validate the effectiveness of the proposed hybrid forecasting architecture, a number of comparable experiments are provided***. Besides the hybrid FEEMD-CG-BA-RBFNN model, the hybrid FEEMD-CG-BA-GRNN model, the hybrid SSA-CG-BA-RBFNN model and the hybrid SSA-CG-BA-GRNN model, the single RBFNN model, the single GRNN model and the single ARIMA (autoregressive integrated moving average) model are also included in the performance comparison to obtain the best combination in the proposed architecture.

The remainder of the paper is organized as follows. Section 2 introduces the hybrid forecasting strategy proposed in this paper. Section 3 presents the wind speed decomposition contained in the hybrid forecasting strategy. Section 4 develops a new improved optimization algorithm. Section 5 proposes four hybrid models. The forecasting results of the proposed hybrid models and comparisons are discussed in Section 6. Finally, Section 7 concludes the paper.

2. Framework of the proposed hybrid architecture

The flowchart of the proposed hybrid architecture in this study is given in Fig. 2. In Fig. 2, the proposed study can be summarized briefly as follows:

- Decompose the original wind speed time series with the FEEMD algorithm and the SSA algorithm into several sub-layers.
- Build the modified neural networks, CG-BA-RBFNN and CG-BA-GRNN to predict each wind speed sub-layers for the one-step, two-step and three-step prediction.
- Summarize the one-step, two-step and three-step predicted results of each sub-layers from FEEMD and SSA to obtain the final results of CG-BA-RBFNN and CG-BA-GRNN.
- Compare the forecasting performance of each model and find the best one. The compared algorithms include four hybrid models, i.e. FEEMD-CG-BA-RBFNN, FEEMD-CG-BA-GRNN, SSA-CG-BA-RBFNN and SSA-CG-BA-GRNN, in the proposed architecture, four comparison hybrid models, i.e. FEEMD-BA-RBFNN, FEEMD-BA-GRNN, SSA-BA-RBFNN and SSA-BA-GRNN and three single models, i.e. RBFNN, GRNN and ARIMA.

3. Wind speed decomposition

In this paper, two decomposing methods, the SSA algorithm and the FEEMD algorithm, are employed to process the original wind speed data. More information about SSA and FEEMD are shown in Appendix A.

4. CG-BA

In this part, CG-BA is proposed and four test functions are employed to evaluate this developed algorithm.

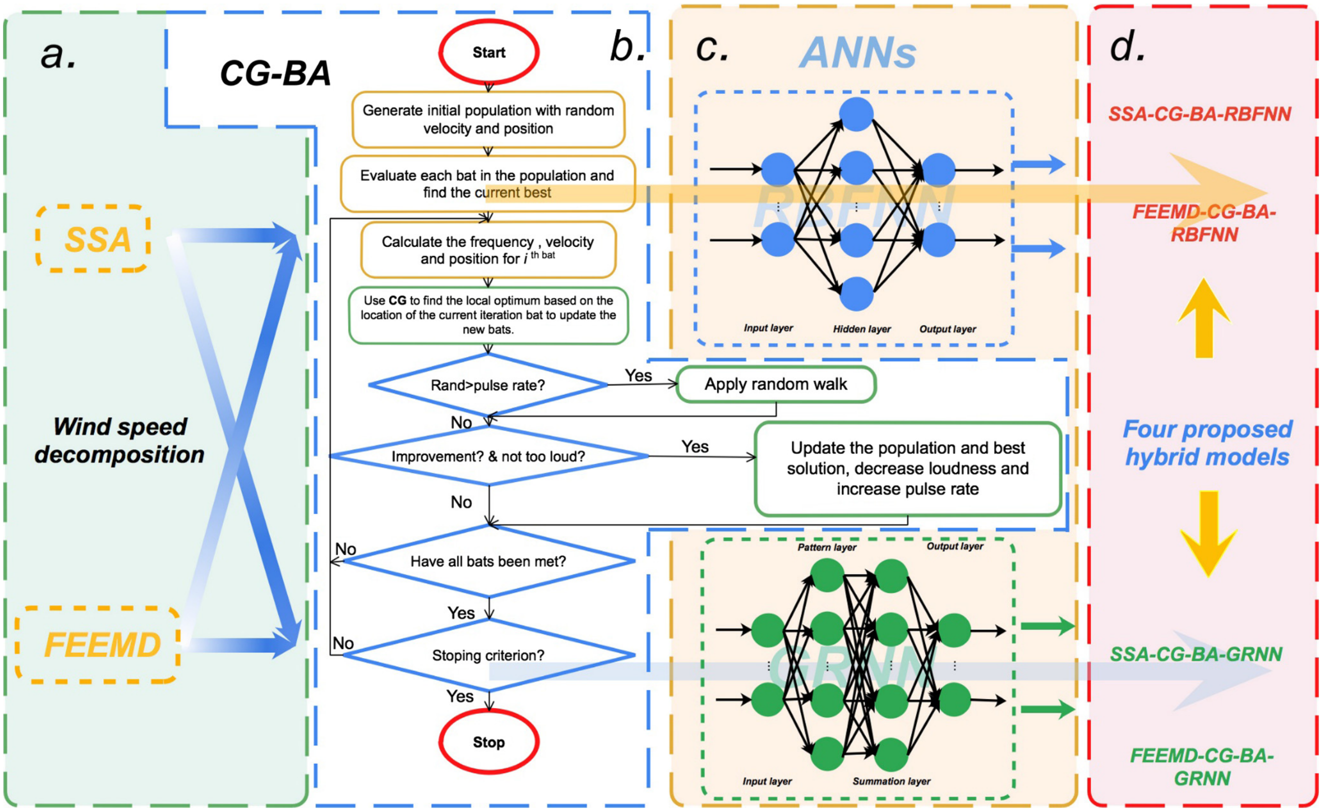


Fig. 2. Structure of hybrid forecasting strategy.

Table 1
Test functions.

Function name	Modal characteristic	Test function	Variable domain	Global optimum
Sphere	Unimodal	$f(x) = \sum_{i=1}^{d} x_i^2$	$x_i \in [-5.12, 5.12]$	$f_{\min}(0,0,0\cdots0) = 0$
Rosenbrock	Multi-modal	$f(X) = \sum_{i=1}^{d-1} [100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2]$	$x_i \in [-2.084, 2.084]$	$f_{\min}(1,1,1\cdots1) = 0$
Rastrigin	Multi-modal	$f(X) = \sum_{i=1}^{d} (x_i^2 - 10(2\pi x_i) + 10)$	$x_i \in [-5.12, 5.12]$	$f_{\min}(0,0,0\cdots0) = 0$
Schaffer	Multi-modal	$f(X) = \frac{\sin^2\sqrt{\sum_{i=1}^{d} x_i^2} - 0.5}{\left[1 + 0.001\left(\sum_{i=1}^{d} x_i^2\right)\right]^2} + 0.5$	$x_i \in [-5.12, 5.12]$	$f_{\min}(0,0,0\cdots0) = 0$

4.1. CG-BA

The bat algorithm is a novel optimization algorithm proposed by Yang [48,49], which was inspired by the echolocation behavior of natural bats in determining their foods. BA not only offers powerful global exploration and exploitation abilities but also has the good ability to find the local optimum.

However, conventional BA continues to suffer from slow convergence during the later period of optimization when it is applied to large-scale and complex problems. To speed up the convergence, a new improved BA based on the CG quasi-Newton method was developed in this work. CG is developed on the basis of the Newton algorithm and the steepest descent algorithm (SDA). Meanwhile, the shortcomings of these two algorithms, slow convergence of SDA and complex computation of Newton algorithm, are overcome with CG [50]. As shown in Fig. 2 Part b, it is used when BA updates solutions in an iteration to find a local optimal solution and thus enhance the local optimization ability and the speed of the local convergence of the whole algorithm.

Let $\boldsymbol{x}^{iter}$ be the positions of bats in BA, where *iter* represents the current iteration number. Generally, $\boldsymbol{x}^{iter}$ is input into the fitness function directly to evaluate the current best value. To improve the local search ability of BA, a CG circulation is added in BA. In this circulation, $\boldsymbol{x}^{iter}$ will be the initial value to continue searching with the gradient $-\nabla f(\boldsymbol{x}_j^{iter})$ and step length λ_j^{iter}. This iterative loop could be presented as

$$\boldsymbol{x}_{j+1}^{iter} = \boldsymbol{x}_j^{iter} + \lambda_j^{iter}\boldsymbol{d}_j^{iter}, \quad (j = 0, 1, \ldots, M-1) \tag{1}$$

where $\boldsymbol{d}_j^{iter} = -\nabla f(\boldsymbol{x}_j^{iter})$, and M is the total number of CG iterations.

After processing the CG circulation, a new position of $\boldsymbol{x}_M^{iter}$ is obtained. $\boldsymbol{x}_M^{iter}$ is input into the fitness function to evaluate a value as the current best result. Then $\boldsymbol{x}_M^{iter}$ is updated to $\boldsymbol{x}_0^{iter+1}$ according to the BA rules [48,49]. Meanwhile, to keep the results from trapping in local optimums, a lot of experiments have been done to select the total iteration number M. Finally we find that when M is set in the region from 4 to 6, the optimization performance is good. If M is less than 4, the rate of convergence may not be enhanced. And if M is greater than 6, the results are easily trapped in local optimums. The pseudo code for CG-BA is provided in Algorithm 1.

Algorithm 1. CG-BA

Output:
$\boldsymbol{x}_b$—the value of $\boldsymbol{x}$ with the best fitness value in the population

Parameters:
α—a random vector, with a value between 0 and 1.
ε—a random vector, with a value between 0 and 1.
β—a random vector, with a value between 0 and 1.
F_i—the fitness function of $\boldsymbol{x}_i$.
N—the number of generations P.
t—current iteration number.
$Iter_{\mathbf{max}}$—the maximum number of iterations.
M—the maximum number of iterations of conjugate gradient algorithm.
L—the loudness of a bat.
r—the pulse rate of a bat.
$\boldsymbol{v}_i^t$ —the velocity of $\boldsymbol{x}_i$ at iteration t.
$\boldsymbol{x}_i^t$ —the position of $\boldsymbol{x}_i$ at iteration t.
$\boldsymbol{g}_i^k$ —the gradient of $\boldsymbol{x}_i$ at iteration k.
$\boldsymbol{d}_i^k$ —the search direction of $\boldsymbol{x}_i$ at iteration k.
λ^k —the steplength of $\boldsymbol{x}_i$ at iteration k.

/*Initialize generation P ($\boldsymbol{x}_i$,i=1,2,…,N) in random positions.*/
/*Initialize t=0.*/
FOR EACH i=1:N **DO**
  Evaluate the corresponding fitness function F_i
END FOR
WHILE $t<Iter_{\max}$ **DO**
  FOR EACH i=1:N **DO**
    $F_i=F_{\min}+(F_{\max}-F_{\min})\alpha$
    $\boldsymbol{v}_i^t=\boldsymbol{v}_i^{t-1}+\left(\boldsymbol{x}_i^t-\boldsymbol{x}^*\right)F_i$
    $\boldsymbol{x}_i^t=\boldsymbol{x}_i^{t-1}+\boldsymbol{v}_i^t$
  END FOR
  /*Use conjugate gradient algorithm.*/
  FOR EACH k=1:M **DO**
    FOR EACH i=1:N **DO**
      $\boldsymbol{g}_i^k=\nabla F\left(\boldsymbol{x}_i^k\right)$
      $\varphi^{k-1}=\dfrac{\left\|\boldsymbol{g}^k\right\|^2}{\left\|\boldsymbol{g}^{k-1}\right\|^2}$
      $\boldsymbol{d}_i^k=-\boldsymbol{g}_i^k+\varphi^{k-1}\boldsymbol{d}_i^{k-1}$
      $\lambda^k=-\dfrac{\left(\boldsymbol{g}^k\right)^T\boldsymbol{d}^k}{\left(\boldsymbol{d}^k\right)^T A\boldsymbol{d}^k}$
      /*Where A is the symmetric positive definite matrix.*/
      $\boldsymbol{x}_i^{k+1}=\boldsymbol{x}_i^k+\lambda^k\boldsymbol{d}_i^k$
    END FOR
  END FOR
  /*Update the current best solution $\boldsymbol{x}^*$.*/
  FOR EACH i=1:N **DO**
    Evaluate the corresponding fitness function F_i
  END FOR
  IF $F_{best}<F^*$ **THEN**
    FOR EACH i=1:N **DO**
      $\boldsymbol{x}_{new}=\boldsymbol{x}_{old}+\varepsilon L^t$
      IF $L^t>\beta$ **THEN**
        $L_i^{t+1}=\beta L_i^t$
        $r_i^{t+1}=r_i^0\left[1-\exp\left(-\gamma t\right)\right]$
      END IF
    END FOR
  END IF
  $iter=iter+1$
END WHILE
RETURN $\boldsymbol{x}_{best}$

Table 2
The experimental parameters of BA and CG-BA.

Experimental parameters	BA	CG-BA
Maximum generation	10,000	10,000
Population size	100	100
Convergence tolerance	10^{-5}	10^{-5}
Maximum generation of CG	–	5

4.2. Test of CG-BA

To evaluate the proposed algorithm, CG-BA, four test functions are employed as shown in Table 1. Sphere function is unimodal, Rosenbrock's function is multimodal, Rastrigin's function is multimodal and Schaffer function is multimodal. The tests of BA and CG-BA on all test functions were performed on an Intel i7-4870 2.50 GHz machine with 16 GB RAM. The experimental parameters of BA and CG-BA are shown in Table 2.

As the test results shown in Table 3, two points can be concluded:

(a) The max value, min value and average value of iteration of CG-BA are less than the original BA for four test functions. This means the convergence ability of BA has been successfully improved with CG-BA.
(b) For the Rosenbrock's function, Rastrigin's function and Schaffer function, the convergence rates of BA didn't obtain 1. While for these three test functions, the convergence rates of CG-BA are obtained 1. Thus the optimization performance of the original BA has also enhanced with CG-BA.

Remark. Through the experimental results and above analysis, the optimization ability of the original BA has been successfully enhanced by the proposed CG-BA.

5. Optimization of RBFNN and GRNN

The proposed CG-BA is employed to optimize the initial weights and thresholds for the RBFNN and GRNN.

Table 5
Experimental parameter setting.

Model	Experimental parameters	Default value
GRNN	Neuron number of the input layer	4
	Neuron number of the hidden layer	9
	Neuron number of the output layer	1
	Radial basis function expansion	0.1 to 2.0
	Maximum number of training	1000
	Training requirement precision	0.00002
RBFNN	Neuron number of the input layer	4
	Neuron number of the hidden layer	9
	Neuron number of the output layer	1
	Sample	400
	Maximum number of training	1000
	Training requirement precision	0.00002

5.1. RBFNN optimized by CG-BA

This section contains the standard RBFNN and the improved RBFNN that is optimized by CG-BA.

5.1.1. Standard RBFNN

The structure of RBFNN is simple and includes an input layer, a hidden layer and an output layer, as shown in Fig. 2 Part c. The hidden layer is the key part of RBFNN, and its neurons represent RBFNN. More information about RBFNN is shown in Appendix B.

5.1.2. RBFNN optimized by CG-BA

The final results are dependent on the initial random weights and threshold values of an ANN, which will increase the unstable factor in forecasting. In this part of the paper, the CG-BA-RBFNN model is developed, and the proposed optimization algorithm CG-BA is used to optimize the initial weight and threshold of the RBFNN to improve the forecasting performance of RBFNN. The details of the CG-BA-RBFNN is presented as Algorithm 2.

Table 3
Test results of BA and CG-BA.

Test function	Dimension	Algorithm	Max value of iteration	Min value of iteration	Average value of iteration	Convergence rate
Sphere	10	BA	213	174	198	1
		CG-BA	3	1	1.4	1
	20	BA	343	151	182	1
		CG-BA	21	12	17.2	1
	50	BA	513	398	441	1
		CG-BA	112	84	98	1
Rosenbrock	2	BA	–	–	–	–
		CG-BA	165	79	103	1
Rastrigin	10	BA	421	315	369	1
		CG-BA	197	144	182	1
	20	BA	860	731	795	0.83
		CG-BA	528	378	469	1
	50	BA	–	–	–	–
		CG-BA	1342	873	1128	0.96
Schaffer	2	BA	1236	981	1035	0.81
		CG-BA	56	24	43	1

Table 4
Statistical parameters for the data used in this paper.

Region	Mean value (m/s)	Std. dev. (m/s)	Maximum value (m/s)	Minimum value (m/s)	Median value (m/s)
Site 1	6.5564	2.4147	14.3000	1.4000	6.3000
Site 2	5.8237	2.1162	15.7000	0.9000	5.6000
Site 3	7.8363	3.4319	18.3000	1.0000	7.1000
Site 4	6.4602	2.5402	17.2000	0.8000	6.3000

Algorithm 2. CG-BA-RBFNN

Input:

$x_t^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \mathrm{K}, x^{(0)}(q)\right)$–sequence of training wind speed data.

$x_v^{(0)} = \left(x^{(0)}(q+1), x^{(0)}(q+2), \mathrm{K}, x^{(0)}(q+d)\right)$–sequence of verification wind speed data

Output:

$\hat{y}_z^{(0)} = \left(\hat{y}^{(0)}(q+1), \hat{y}^{(0)}(q+2), \mathrm{K}, \hat{y}^{(0)}(q+d)\right)$—the forecasting electrical load data from RBFNN

Parameters:

α—a random vector, with a value between 0 and 1.
ε—a random vector, with a value between 0 and 1.
β—a random vector, with a value between 0 and 1.
F_i—the fitness function of x_i.
N—the number of generations P.
t—current iteration number.
$Iter_{\mathbf{max}}$—the maximum number of iterations.
M—the maximum number of iterations of conjugate gradient algorithm.
L—the loudness of a bat.
r—the pulse rate of a bat.
$\boldsymbol{x}_i$—generation i (the weight and threshold of the RBFNN)

Fitness function

$$F = \sum_{i=1}^{d} \left| \hat{x}_v^{(0)}(q+i) - \hat{y}_z^{(0)}(q+d) \right|$$

/*Initialize generation $\boldsymbol{X}$ ($\boldsymbol{x}_i$,i=1, 2,…, N, the weight and threshold of the RBFNN) in random positions.*/
/*Initialize t=0.*/
FOR EACH i=1:N **DO**
  Evaluate the corresponding fitness function F_i
END FOR
WHILE t<$Iter_{\max}$ **DO**
  FOR EACH i=1:N **DO**
    F_i=$F_{\min}$+($F_{\max}$-$F_{\min}$)α
    $\boldsymbol{v}_i^t = \boldsymbol{v}_i^{t-1} + \left(\boldsymbol{x}_i^t - \boldsymbol{x}^*\right) F_i$
    $\boldsymbol{x}_i^t = \boldsymbol{x}_i^{t-1} + \boldsymbol{v}_i^t$
  END FOR

/*Use conjugate gradient algorithm.*/
FOR EACH k=1:M **DO**
  FOR EACH i=1:N **DO**
    $g_i^k = \nabla F\left(x_i^k\right)$
    $\varphi^{k-1} = \dfrac{\left\|g^k\right\|^2}{\left\|g^{k-1}\right\|^2}$
    $d_i^k = -g_i^k + \varphi^{k-1} d_i^{k-1}$
    $\lambda^k = -\dfrac{\left(g^k\right)^T d^k}{\left(d^k\right)^T A d^k}$
    /*Where A is the symmetric positive definite matrix.*/
    $x_i^{k+1} = x_i^k + \lambda^k d_i^k$
  END FOR
END FOR
/*Update the current best solution x^*.*/
FOR EACH i=1:N **DO**
  Evaluate the corresponding fitness function F_i
END FOR
IF $F_{best} < F^*$ **THEN**.
  FOR EACH i=1:N **DO**
    $x_{new}=x_{old}+\varepsilon L^t$
    IF $L^t>$ **THEN**
      $L_i^{t+1} = \beta L_i^t$
      $r_i^{t+1} = r_i^0\left[1-\exp(-\gamma t)\right]$
    END IF
  END FOR
END IF
$iter = iter + 1$
END WHILE
RETURN x_b
Set the weight and threshold of the RBFNN according to x_b.
Use x_t to train the RBFNN and update the weight and threshold of the RBFNN.
Input the historical data into RBFNN to obtain the forecasting value $\hat{y}$.

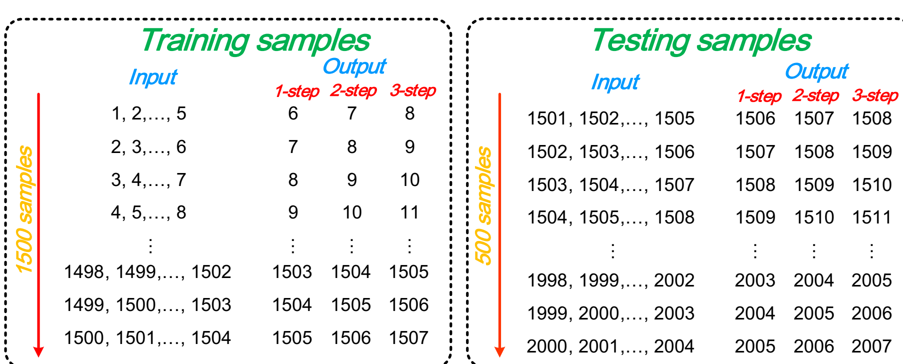


Fig. 3. Input and output data selection.

5.2. GRNN optimized by CG-BA

This section contains the standard GRNN and the improved GRNN that is optimized by CG-BA.

5.2.1. Standard GRNN

Specht [51] proposed a new type of neural network model named GRNN, which is based on the advantage of a standard statistical technology known as Kernel regression [52,53]. Four layers, i.e., the input layer, pattern layer, summation layer, and output layer, compose a GRNN, as shown in Fig. 2 Part c. More information about GRNN is shown in Appendix B.

5.2.2. GRNN optimized by CG-BA

This part proposed that the initial weight and threshold of the GRNN is optimized by the proposed optimization algorithm

CG-BA to improve the forecasting performance. The details of the CG-BA-GRNN are presented as Algorithm 3.

Algorithm 3. CG-BA-GRNN

Input:

$x_t^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \mathrm{K}, x^{(0)}(q)\right)$ –sequence of training wind speed data.

$x_v^{(0)} = \left(x^{(0)}(q+1), x^{(0)}(q+2), \mathrm{K}, x^{(0)}(q+d)\right)$ –sequence of verification wind speed data

Output:

$\hat{y}_z^{(0)} = \left(\hat{y}^{(0)}(q+1), \hat{y}^{(0)}(q+2), \mathrm{K}, \hat{y}^{(0)}(q+d)\right)$ —the forecasting wind speed from GRNN

Parameters:

α—a random vector, with a value between 0 and 1.
ε—a random vector, with a value between 0 and 1.
β—a random vector, with a value between 0 and 1.
F_i—the fitness function of x_i.
N—the number of generations P.
t—current iteration number.
$Iter_{\mathbf{max}}$—the maximum number of iterations.
M—the maximum number of iterations of conjugate gradient algorithm.
L—the loudness of a bat.
r—the pulse rate of a bat.
$\boldsymbol{x}_i$—generation i (the weight and threshold of the GRNN)

Fitness function

$$F = \sum_{i=1}^{d} \left| \hat{x}_v^{(0)}(q+i) - \hat{y}_z^{(0)}(q+d) \right|$$

/*Initialize generation $\boldsymbol{X}(\boldsymbol{x}_i, i=1,2,\ldots,N,$ the weight and threshold of the GRNN) in random positions.*/
/*Initialize t=0.*/
FOR EACH i=1:N **DO**
    Evaluate the corresponding fitness function F_i
END FOR
WHILE $t<Iter_{\max}$ **DO**
    FOR EACH i=1:N **DO**
        $F_i=F_{\min}+(F_{\max}-F_{\min})\alpha$
        $\boldsymbol{v}_i^t = \boldsymbol{v}_i^{t-1} + \left(\boldsymbol{x}_i^t - \boldsymbol{x}^*\right) F_i$
        $\boldsymbol{x}_i^t = \boldsymbol{x}_i^{t-1} + \boldsymbol{v}_i^t$
    END FOR
    /*Use conjugate gradient algorithm.*/

FOR EACH k=1:M **DO**
  FOR EACH i=1:N **DO**
    $\boldsymbol{g}_i^k = \nabla F\left(\boldsymbol{x}_i^k\right)$
    $\varphi^{k-1} = \dfrac{\left\|\boldsymbol{g}^k\right\|^2}{\left\|\boldsymbol{g}^{k-1}\right\|^2}$
    $\boldsymbol{d}_i^k = -\boldsymbol{g}_i^k + \varphi^{k-1}\boldsymbol{d}_i^{k-1}$
    $\lambda^k = -\dfrac{\left(\boldsymbol{g}^k\right)^T \boldsymbol{d}^k}{\left(\boldsymbol{d}^k\right)^T A\boldsymbol{d}^k}$
    /*Where A is the symmetric positive definite matrix.*/
    $\boldsymbol{x}_i^{k+1} = \boldsymbol{x}_i^k + \lambda^k \boldsymbol{d}_i^k$
  END FOR
END FOR
/*Update the current best solution $\boldsymbol{x}^*$.*/
FOR EACH i=1:N **DO**
  Evaluate the corresponding fitness function F_i
END FOR
IF $F_{best} < F^*$ **THEN**.
  FOR EACH i=1:N **DO**
    $\boldsymbol{x}_{new} = \boldsymbol{x}_{old} + \varepsilon L^t$
    IF $L^t >$ **THEN**
      $L_i^{t+1} = \beta L_i^t$
      $r_i^{t+1} = r_i^0\left[1-\exp(-\gamma t)\right]$
    END IF
  END FOR
END IF
$iter = iter + 1$
END WHILE
RETURN $\boldsymbol{x}_b$
Set the weight and threshold of the GRNN according to $\boldsymbol{x}_b$.
Use $\boldsymbol{x}_t$ to train the GRNN and update the weight and threshold of the GRNN.
Input the historical data into GRNN to obtain the forecasting value $\hat{y}$.

6. Forecasting experiment

In this part, the experiments were divided into three parts, ***Experiment I***, ***Experiment II*** and ***Experiment III***. In ***Experiment I***, the wind speed multi-step forecasting results and comparisons of the hybrid FEEMD-CG-BA-RBFNN model, the hybrid FEEMD-CG-BA-GRNN model, the hybrid SSA-CG-BA-RBFNN model and the hybrid SSA-CG-BA-GRNN model are given. In ***Experiment II***, the performance of the hybrid FEEMD-CG-BA-RBFNN model, the hybrid FEEMD-CG-BA-GRNN model, the hybrid SSA-CG-BA-RBFNN model and the hybrid SSA-CG-BA-GRNN model are compared with FEEMD-BA-RBFNN, FEEMD-BA-GRNN, SSA-BA-RBFNN, SSA-BA-GRNN, single RBFNN, single GRNN and ARIMA. In ***Experiment III***, the DM-test is used to evaluate the performance of each forecasting model. To confirm the universality of the proposed model, ***Experiment I***, ***Experiment II*** and ***Experiment III*** are validated at four different sites.

Four data sites in Penglai region have been selected with the latitude from 120°43′N to 120°47′N and longitude from 37°50′E to 37°37′E. The data sites are in a mountain and hilly area near the sea, and its altitude ranges from 100 m to 240 m. The rated power of WTG (wind power generator) is 1500 kW. The mean annual temperature, humidity and air pressure in this region are 11.9 °C, 65% and 1012.7 hPa, respectively. Statistical parameters for the data used in this paper are shown in Table 4.

All algorithms are operated on the following platform: MATLAB R2012a on Windows 8 with 2.50 GHz Intel Core i7 4870HQ 64-bit and 16 GB of RAM. The experimental parameters are shown in Table 5. Meanwhile, considering randomness factors and ensuring that the final results are reliable and independent of the initial weights, we carry out each experiment 50 times and then take the average value. The input layer of all the ANNs is constructed with four neurons. Hecht–Nelson method [54] is employed to determine the node number of the hidden layer. When the node number of the input layer is n, the node number of the hidden layer is $2n + 1$.

6.1. Accuracy estimating indexes

To learn the global traits of the models, three metric parameters are taken: the MAE (mean absolute error), the MAPE (mean absolute percentage error) and the MSE (mean square error). MAE is the average absolute forecast error of n times forecast results. Because the prediction error may be positive and negative, it cannot reflect the level of error; this problem can be avoided by using MAE. MSE is the average of the prediction error squares, which can

Table 6
Forecasting performance of four proposed hybrid models.

		SSA-CG-BA-RBFNN			SSA-CG-BA-GRNN			FEEMD-CG-BA-RBFNN			FEEMD-CG-BA-GRNN		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	MAE	0.0906	0.1783	0.2799	0.0698	0.1621	0.2334	0.1047	0.1869	0.2963	0.0755	0.1661	0.2581
	MSE	0.0091	0.0353	0.0871	0.0054	0.0291	0.0605	0.0122	0.0389	0.0977	0.0063	0.0306	0.0740
	MAPE (%)	1.6510	3.2489	5.1002	1.2720	2.9528	4.2507	1.9053	3.4029	5.3953	1.3752	3.0261	4.7022
Site 2	MAE	0.1011	0.2155	0.3121	0.0773	0.1767	0.2483	0.0948	0.2024	0.3379	0.0862	0.1881	0.2861
	MSE	0.0122	0.0555	0.1163	0.0071	0.0373	0.0738	0.0107	0.0489	0.1364	0.0089	0.0422	0.0977
	MAPE (%)	1.5737	3.3484	4.8506	1.2016	2.7491	3.8576	1.4738	3.1483	5.2544	1.3405	2.9228	4.4539
Site 3	MAE	0.0938	0.2195	0.3241	0.0851	0.1755	0.2587	0.0955	0.2127	0.3472	0.0903	0.2088	0.3121
	MSE	0.0105	0.0574	0.1252	0.0086	0.0367	0.0795	0.0108	0.0538	0.1435	0.0097	0.0519	0.1159
	MAPE (%)	1.3249	3.0974	4.5735	1.2031	2.4749	3.6545	1.3478	3.0035	4.8984	1.2746	2.9477	4.4049
Site 4	MAE	0.1001	0.2193	0.3633	0.0919	0.2031	0.2846	0.1042	0.2251	0.3645	0.0975	0.2215	0.2903
	MSE	0.0114	0.0548	0.1502	0.0096	0.0469	0.0922	0.0123	0.0577	0.1513	0.0108	0.0559	0.0959
	MAPE (%)	1.3448	2.9471	4.8832	1.2347	2.7263	3.8254	1.4008	3.0241	4.8963	1.3099	2.9763	4.3003

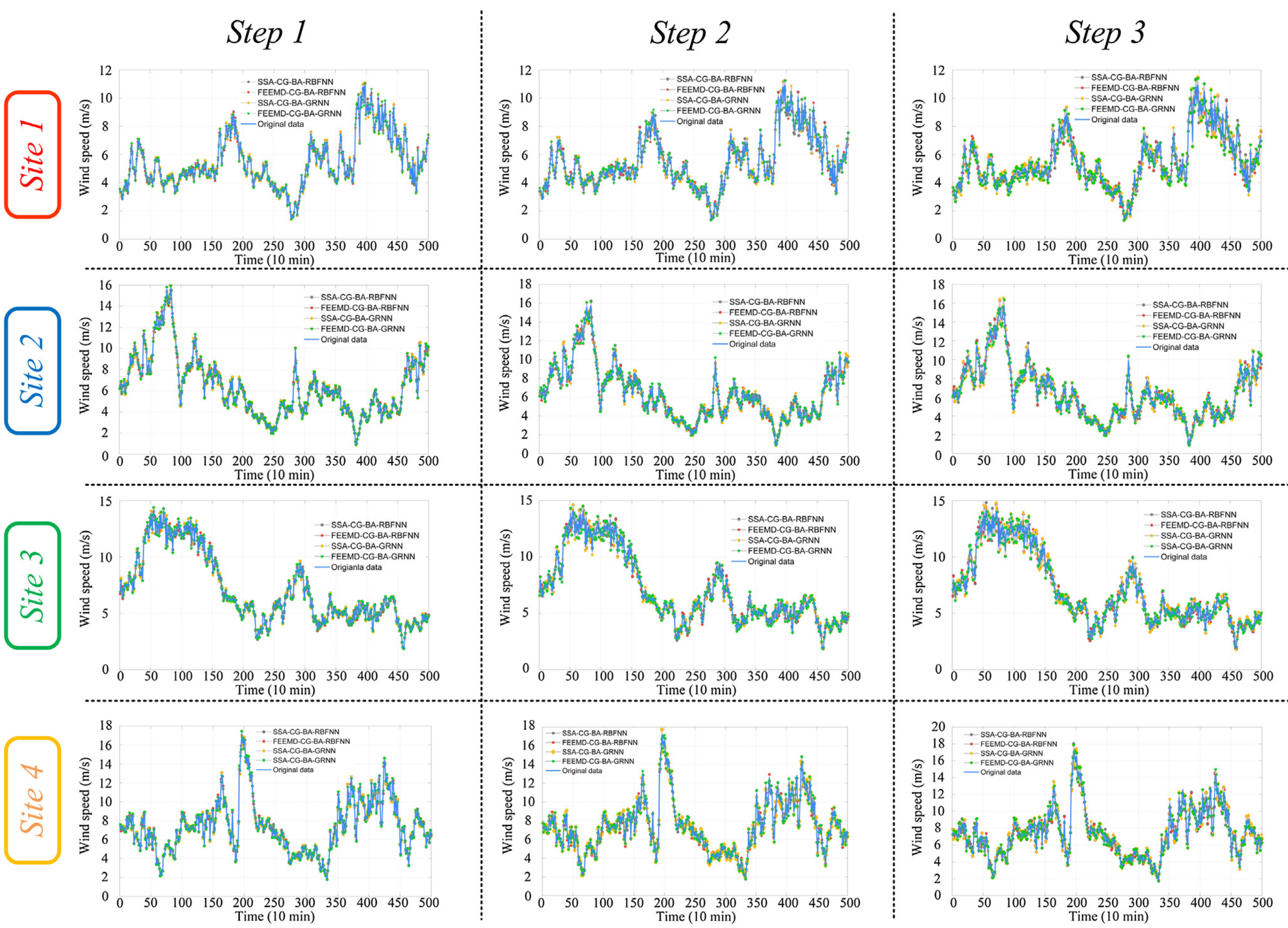


Fig. 4. Forecasting results of four proposed hybrid models.

evaluate the change of the prediction model; the smaller the MSE value is, the better the prediction model is. MAPE is a measure of the accuracy of the prediction method for use in the performance evaluation and comparison in statistics. The detailed equations of these three error indexes are given in Appendix C.

6.2. Experiment I

The data from four wind power stations in Penglai, China are used as test data in this experiment, we chose 1500 history data for training and 500 data for testing as shown in Fig. 3. The multi-step forecasting results of SSA-CG-BA-RBFNN, SSA-CG-BA-GRNN, FEEMD-CG-BA-RBFNN and FEEMD-CG-BA-GRNN are shown in Table 6 and Fig. 4. The detailed multi-step promoting percentages of the hybrid models of the four sites are shown in Table 7 and Fig. 5.

Table 6 and Fig. 4 indicate the following:

(a) For Site 1, when the forecasting is 1-step, SSA-CG-BA-GRNN has the highest accuracy forecasting results with a 1.2720% MAPE value. The second-highest to fourth-highest accurate models are FEEMD-CG-BA-GRNN,

Table 7
Improvement percentages of four proposed hybrid models.

		SSA-CG-BA-RBFNN *vs.* SSA-CG-BA-GRNN			SSA-CG-BA-RBFNN *vs.* FEEMD-CG-BA-RBFNN			SSA-CG-BA-RBFNN *vs.* FEEMD-CG-BA-GRNN		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	ξ_{MAE} (%)	−22.958	−9.086	−16.613	15.563	4.823	5.859	−16.667	−6.842	−7.788
	ξ_{MSE} (%)	−40.659	−17.564	−30.540	34.066	10.198	12.170	−30.769	−13.314	−15.040
	ξ_{MAPE} (%)	−22.956	−9.114	−16.656	15.403	4.740	5.786	−16.705	−6.858	−7.804
Site 2	ξ_{MAE} (%)	−23.541	−18.005	−20.442	−6.231	−6.079	8.267	−14.738	−12.715	−8.331
	ξ_{MSE} (%)	−41.803	−32.793	−36.543	−12.295	−11.892	17.283	−27.049	−23.964	−15.993
	ξ_{MAPE} (%)	−23.645	−17.898	−20.472	−6.348	−5.976	8.325	−14.819	−12.711	−8.178
Site 3	ξ_{MAE} (%)	−9.275	−20.046	−20.179	1.812	−3.098	7.127	−3.731	−4.875	−3.703
	ξ_{MSE} (%)	−18.095	−36.063	−36.502	2.857	−6.272	14.617	−7.619	−9.582	−7.428
	ξ_{MAPE} (%)	−9.193	−20.098	−20.094	1.728	−3.032	7.104	−3.797	−4.833	−3.686
Site 4	ξ_{MAE} (%)	−8.192	−7.387	−21.663	4.096	2.645	0.330	−2.597	1.003	−20.094
	ξ_{MSE} (%)	−15.789	−14.416	−38.615	7.895	5.292	0.732	−5.263	2.007	−36.152
	ξ_{MAPE} (%)	−8.187	−7.492	−21.662	4.164	2.613	0.268	−2.595	0.991	−11.937
		SSA-CG-BA-GRNN *vs.* FEEMD-CG-BA-RBFNN			SSA-CG-BA-GRNN *vs.* FEEMD-CG-BA-GRNN			FEEMD-CG-BA-RBFNN *vs.* FEEMD-CG-BA-GRNN		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	ξ_{MAE} (%)	50.000	15.299	26.949	8.166	2.468	10.583	−27.889	−11.129	−12.892
	ξ_{MSE} (%)	125.926	33.677	61.488	16.667	5.155	22.314	−48.361	−21.337	−24.258
	ξ_{MAPE} (%)	49.788	15.243	26.927	8.113	2.482	10.622	−27.822	−11.073	−12.846
Site 2	ξ_{MAE} (%)	22.639	14.544	36.085	11.514	6.452	15.224	−9.072	−7.065	−15.330
	ξ_{MSE} (%)	50.704	31.099	84.824	25.352	13.137	32.385	−16.822	−13.701	−28.372
	ξ_{MAPE} (%)	22.653	14.521	36.209	11.560	6.318	15.458	−9.045	−7.163	−15.235
Site 3	ξ_{MAE} (%)	12.221	21.197	34.210	6.110	18.974	20.642	−5.445	−1.834	−10.109
	ξ_{MSE} (%)	25.581	46.594	80.503	12.791	41.417	45.786	−10.185	−3.532	−19.233
	ξ_{MAPE} (%)	12.027	21.358	34.037	5.943	19.104	20.534	−5.431	−1.858	−10.075
Site 4	ξ_{MAE} (%)	13.384	10.832	28.074	6.094	9.060	2.003	−6.430	−1.599	−20.357
	ξ_{MSE} (%)	28.125	23.028	64.100	12.500	19.190	4.013	−12.195	−3.120	−36.616
	ξ_{MAPE} (%)	13.453	10.923	27.994	6.091	9.170	12.414	−6.489	−1.581	−12.172

SSA-CG-BA-RBFNN and FEEMD-CG-BA-RBFNN with MAPE values of 1.3752%, 1.6510% and 1.9053%, respectively. When the forecasting is 2-step, SSA-CG-BA-GRNN has the most accurate forecasting results with a MAPE value of 2.9528%. According to the MAPE value, FEEMD-CG-BA-GRNN is the second most accurate model, SSA-CG-BA-RBFNN is the third most accurate model and FEEMD-CG-BA-RBFNN is the fourth most accurate model with MAPE values of 3.0261%, 3.2489% and 3.4029%, respectively. When the forecasting is 3-step, SSA-CG-BA-GRNN is still the most accurate forecasting model among the proposed four hybrid models.

(b) For Site 2, SSA-CG-BA-GRNN has the most accurate forecasting results among the 1.2720%, 2.7491% and 3.8576%, respectively. When the forecasting is 2-step, FEEMD-CG-BA-RBFNN is more accurate than SSA-CG-BA-RBFNN. In the three-step forecasting, the precision of the hybrid models is ranked from high to low as SSA-CG-BA-GRNN, FEEMD-CG-BA-GRNN, SSA-CG-BA-RBFNN and FEEMD-CG-BA-RBFNN.

(c) SSA-CG-BA-GRNN is the most accurate model for one-step to three-step forecasting among the four hybrid models in Site 3. The CG-BA-RBFNN with the SSA decomposition algorithm is more precise than CG-BA-RBFNN with the FEEMD algorithm.

(d) SSA-CG-BA-GRNN is still the most accurate forecasting model from one-step forecasting to three-step forecasting among all of the proposed models for the data from Site 4. However, for this site, the two-step forecasting results of SSA-CG-BA-RBFNN are more accurate than FEEMD-CG-BA-GRNN.

Table 7 and Fig. 5 shows the following:

(a) In the one-step predictions, the ξ_{MAPE} (%) value indicates that, from four sites, the MAPE value of SSA-CG-BA-RBFNN, FEEMD-CG-BA-GRNN and FEEMD-CG-BA-RBFNN are decreased with SSA-CG-BA-GRNN with −22.956, −23.645, −9.193 and −8.187; 8.113, 11.560, 5.943 and 6.091; −16.705, −14.819, −3.797 and −2.595, respectively.

(b) In the two-step and three-step predictions, SSA-CG-BA-GRNN also decreases the MAPE value based on SSA-CG-BA-RBFNN, FEEMD-CG-BA-GRNN and FEEMD-CG-BA-RBFNN.

(c) In the one-step predictions, the ξ_{MAPE} (%) value illustrates that from four sites, FEEMD-CG-BA-GRNN decreases 27.822%, 9.045%, 5.431% and 6.489% MAPE values based on FEEMD-CG-BA-RBFNN. In the two-step predictions, FEEMD-CG-BA-GRNN decreases 11.073%, 7.163%, 1.858% and 1.581% MAPE values based on FEEMD-CG-BA-RBFNN. 12.846%, 15.235%, 10.075% and 12.172% MAPE values are decreased with FEEMD-CG-BA-GRNN based on FEEMD-CG-BA-RBFNN in the three-step predictions, respectively.

Remark. By comparing the four proposed hybrid models, the SSA-CG-BA-GRNN hybrid model has the most accurate forecasting results. Comparisons of FEEMD-CG-BA-GRNN with FEEMD-CG-BA-RBFNN and SSA-CG-BA-GRNN with SSA-CG-BA-RBFNN could conclude that the forecasting ability of GRNN is stronger than that of RBFNN. Comparisons of SSA-CG-BA-GRNN with FEEMD-CG-BA-GRNN and SSA-CG-BA-RBFNN with FEEMD-CG-BA-GRNN could reveal that the hybrid models combined with SSA are more accurate than the hybrid models combined with FEEMD.

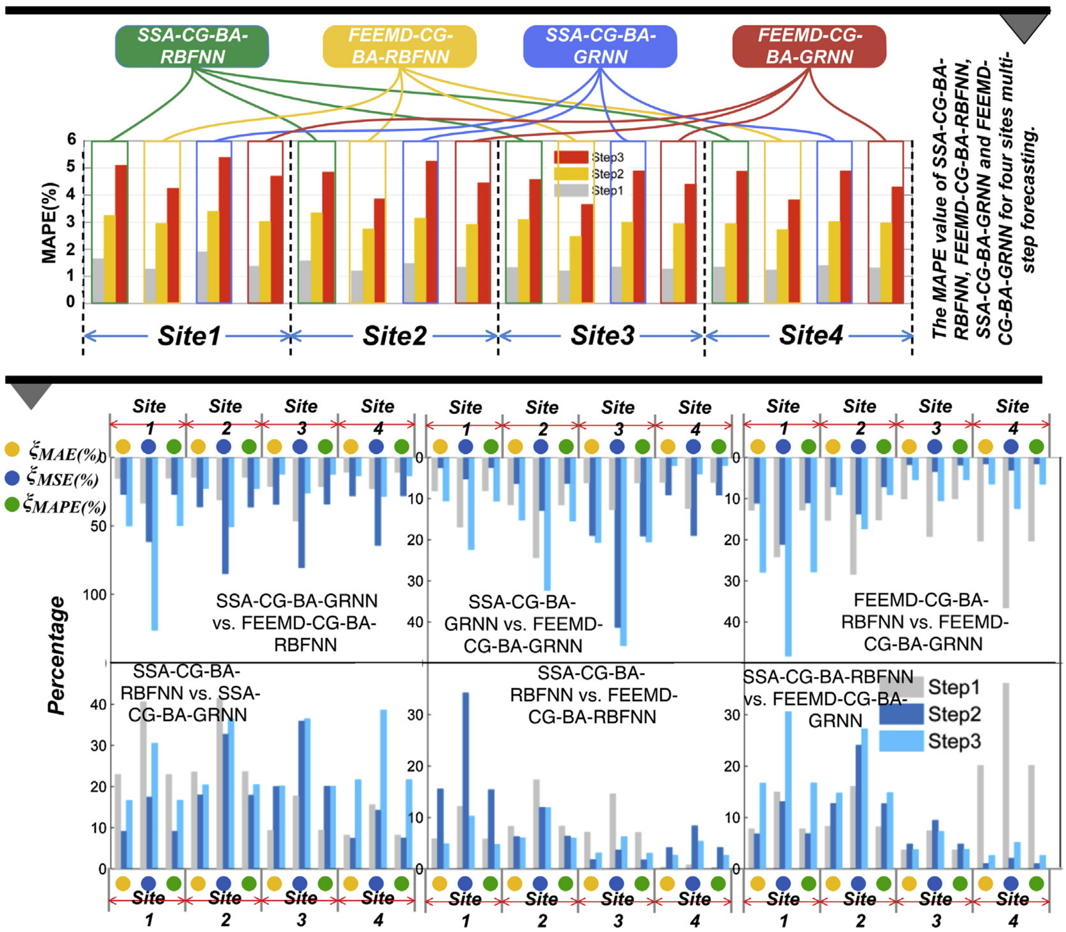


Fig. 5. Promoting percentages of four hybrid models.

Table 8
Forecasting performance of SSA-BA-RBFNN, SSA-BA-GRNN, FEEMD-BA-RBFNN and FEEMD-BA-GRNN.

		SSA-BA-RBFNN			SSA-BA-GRNN			FEEMD-BA-RBFNN			FEEMD-BA-GRNN		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	MAE	0.0971	0.1855	0.2886	0.0825	0.1665	0.2838	0.1628	0.2298	0.3498	0.0962	0.1846	0.4383
	MSE	0.0104	0.0383	0.0926	0.0076	0.0309	0.0728	0.0138	0.0461	0.1034	0.0103	0.0379	0.0732
	MAPE (%)	1.7693	3.3822	5.2634	1.5041	3.0375	4.7621	2.2752	3.7619	5.5838	1.7552	3.3672	5.3457
Site 2	MAE	0.1138	0.2387	0.3335	0.1081	0.1902	0.2638	0.1067	0.2225	0.3333	0.1049	0.2215	0.2961
	MSE	0.0154	0.0581	0.1329	0.0139	0.0432	0.0831	0.0636	0.0892	0.1526	0.0131	0.0586	0.1311
	MAPE (%)	1.7698	3.6455	5.1838	1.6799	2.9566	4.1009	1.7593	3.5575	5.5822	1.6307	3.4444	5.2921
Site 3	MAE	0.1161	0.2403	0.3621	0.1072	0.2197	0.2901	0.1272	0.2446	0.3746	0.1207	0.2438	0.3274
	MSE	0.0161	0.0687	0.1559	0.0137	0.0575	0.1001	0.0193	0.0713	0.1671	0.0174	0.0708	0.1325
	MAPE (%)	1.6391	3.3909	5.1124	1.5125	3.1007	4.0939	1.7951	3.4511	5.2852	1.7036	3.4409	4.7401
Site 4	MAE	0.1273	0.2421	0.3844	0.1221	0.2364	0.3149	0.1385	0.2576	0.3933	0.1304	0.2447	0.3112
	MSE	0.0184	0.0667	0.1683	0.0169	0.0636	0.1128	0.0218	0.0756	0.1761	0.0193	0.0684	0.1303
	MAPE (%)	1.7106	3.2516	5.1642	1.6411	3.1774	4.2304	1.8609	3.4617	5.2849	1.7519	3.2871	4.9806

6.3. Experiment II

This experiment is divided into two parts. The first part illustrates the multi-step forecasting results of SSA-BA-RBFNN, SSA-BA-GRNN, FEEMD-BA-RBFNN and FEEMD-BA-GRNN (as shown in Table 8 and Fig. 6) and the detailed multi-step promoting percentages, using the data from four sites, to evaluate the efficiency of the developed optimization algorithm CG-BA (as shown in Table 9 and Fig. 6). In the second part, five single models, ELM, SVM, RBFNN, GRNN and ARIMA, are employed in multi-step prediction (as shown in Tables 10 and 11). The detailed multi-step promoting percentages of SSA-CG-BA-RBFNN, SSA-CG-BA-GRNN, FEEMD-CG-BA-RBFNN and FEEMD-CG-BA-GRNN by RBFNN, GRNN and ARIMA of four sites are shown in Tables 12 and 13.

Table 8 shows the following:

(a) SSA-BA-GRNN achieves the highest accuracy in one-step prediction to three-step prediction based on the data from four sites.
(b) FEEMD-BA-GRNN is ranked as the second most accurate model among the four models listed in Table 5, except for two-step prediction at Site 4.

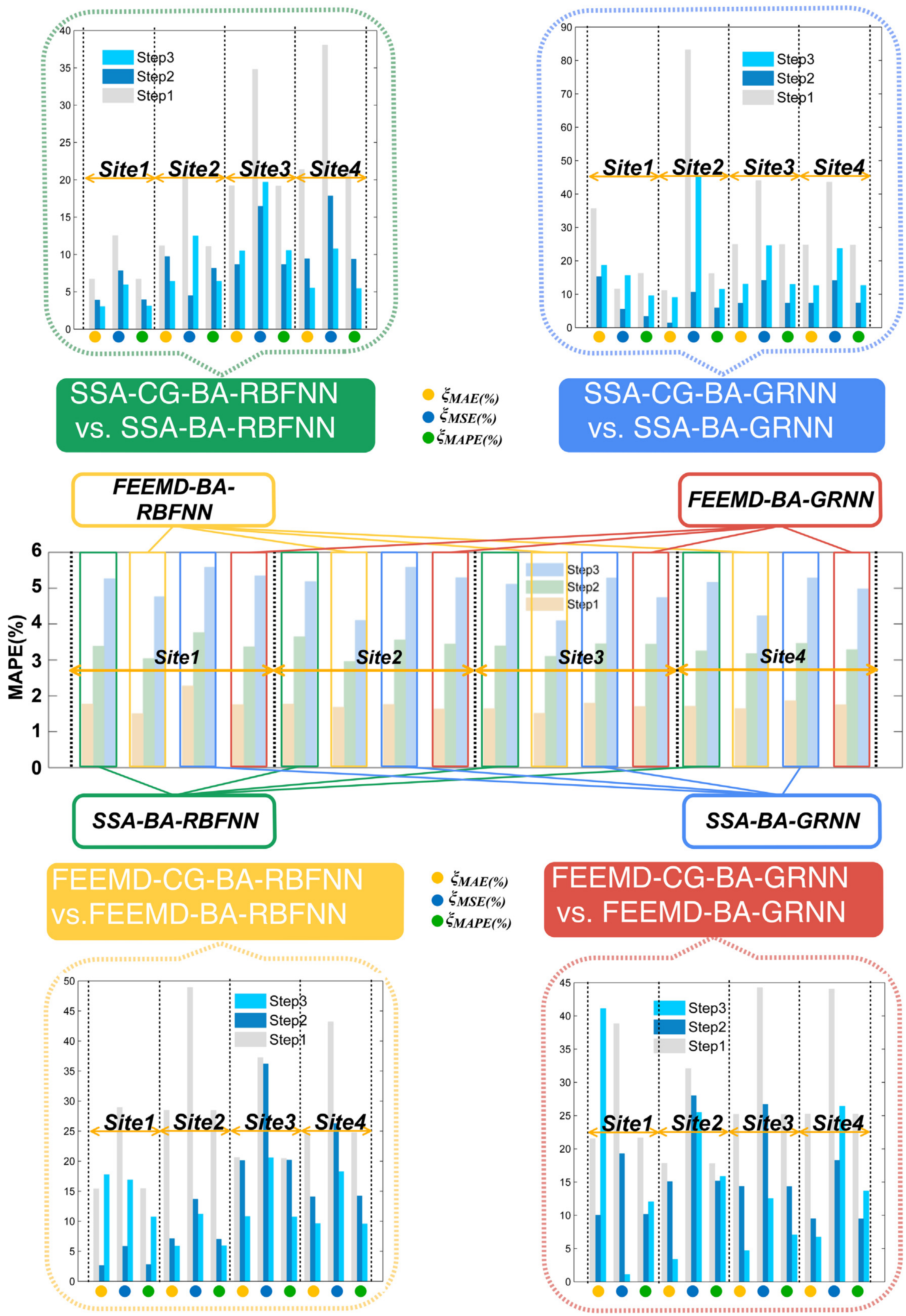


Fig. 6. Forecasting results and promoting percentages of SSA-BA-RBFNN, FEEMD-BA-RBFNN, SSA-BA-GRNN and FEEMD-BA-GRNN.

Table 9
Improvement percentages between four proposed hybrid models and SSA-BA-RBFNN, SSA-BA-GRNN, FEEMD-BA-RBFNN and FEEMD-BA-GRNN.

		SSA-CG-BA-RBFNN *vs.* SSA-BA-RBFNN			SSA-CG-BA-GRNN *vs.* SSA-BA-GRNN			FEEMD-CG-BA-RBFNN *vs.* FEEMD-BA-RBFNN			FEEMD-CG-BA-GRNN *vs.* FEEMD-BA-GRNN		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	ξ_{MAE} (%)	−6.694	−3.881	−3.015	−15.394	−2.643	−17.759	−35.688	−18.668	−15.294	−21.517	−10.021	−41.113
	ξ_{MSE} (%)	−12.534	−7.833	−5.940	−28.947	−5.825	−16.896	−11.594	−15.618	−5.513	−38.834	−19.261	−1.0928
	ξ_{MAPE} (%)	−6.686	−3.941	−3.101	−15.431	−2.788	−10.739	−16.257	−9.543	−3.376	−21.649	−10.131	−12.037
Site 2	ξ_{MAE} (%)	−11.159	−9.719	−6.417	−28.492	−7.098	−5.876	−11.153	−9.034	−1.380	−17.826	−15.079	−3.3772
	ξ_{MSE} (%)	−20.779	−4.475	−12.491	−48.921	−13.657	−11.191	−83.176	−45.179	−10.616	−32.061	−27.986	−25.476
	ξ_{MAPE} (%)	−11.081	−8.150	−6.428	−28.472	−7.018	−5.933	−16.228	−11.503	−5.872	−17.796	−15.143	−15.838
Site 3	ξ_{MAE} (%)	−19.207	−8.656	−10.494	−20.615	−20.118	−10.823	−24.921	−13.041	−7.315	−25.186	−14.356	−4.6731
	ξ_{MSE} (%)	−34.782	−16.448	−19.692	−37.226	−36.173	−20.579	−44.041	−24.544	−14.123	−44.253	−26.694	−12.528
	ξ_{MAPE} (%)	−19.169	−8.656	−10.541	−20.456	−20.182	−10.733	−24.917	−12.969	−7.319	−25.182	−14.333	−7.0716
Site 4	ξ_{MAE} (%)	−21.366	−9.418	−5.489	−24.733	−14.086	−9.622	−24.763	−12.616	−7.323	−25.231	−9.481	−6.7159
	ξ_{MSE} (%)	−38.043	−17.841	−10.754	−43.195	−26.258	−18.262	−43.578	−23.677	−14.082	−44.041	−18.274	−26.401
	ξ_{MAPE} (%)	−21.384	−9.365	−5.441	−24.763	−14.197	−9.574	−24.724	−12.641	−7.353	−25.229	−9.455	−13.659

Table 10
Forecasting performance of RBFNN, GRNN and ARIMA.

		RBFNN			GRNN			ARIMA		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	MAE	0.2523	0.3171	0.4377	0.1861	0.2947	0.3979	0.2912	0.4418	0.6071
	MSE	0.0707	0.1117	0.2131	0.0385	0.0966	0.1761	0.0943	0.2169	0.4095
	MAPE (%)	4.6003	5.7811	7.9794	3.3943	5.3733	7.2559	5.3109	8.0568	11.068
Site 2	MAE	0.2938	0.3846	0.5073	0.2229	0.3395	0.4707	0.3431	0.5241	0.7085
	MSE	0.1031	0.1767	0.3073	0.0594	0.1377	0.2646	0.1405	0.3279	0.5996
	MAPE (%)	4.5687	5.9792	7.8871	3.4642	5.2785	7.3189	5.3339	8.1481	11.013
Site 3	MAE	0.3245	0.4228	0.5418	0.2424	0.3901	0.4853	0.3981	0.5769	0.7797
	MSE	0.1253	0.2128	0.3494	0.0699	0.1811	0.2804	0.1886	0.3961	0.7233
	MAPE (%)	4.5794	5.9651	7.6461	3.4238	5.5041	6.8491	5.6167	8.1408	11.006
Site 4	MAE	0.3223	0.4522	0.5823	0.2571	0.4079	0.5157	0.4049	0.5858	0.8383
	MSE	0.1183	0.2328	0.3861	0.0753	0.1894	0.3028	0.1867	0.3906	0.8006
	MAPE (%)	4.3311	6.0751	7.8223	3.4524	5.4806	6.9291	5.4401	7.8699	11.261

Table 11
Forecasting performance of ELM and SVM.

		ELM			SVM		
		1-step	2-step	3-step	1-step	2-step	3-step
Site 1	MAE	0.1774	0.2756	0.4566	0.2614	0.3015	0.4118
	MSE	0.0381	0.0905	0.1645	0.0495	0.1166	0.1849
	MAPE (%)	3.1044	5.0019	6.9985	3.6498	5.7654	7.6811
Site 2	MAE	0.1956	0.3141	0.4415	0.2415	0.3124	0.4845
	MSE	0.0428	0.1124	0.2561	0.0469	0.1244	0.3146
	MAPE (%)	3.2107	5.0163	6.9875	3.7105	5.8647	7.6447
Site 3	MAE	0.2014	0.3421	0.4685	0.2768	0.3216	0.4975
	MSE	0.0419	0.1031	0.2541	0.0684	0.1467	0.3017
	MAPE (%)	3.1964	5.1651	6.9541	3.8004	5.9451	7.1847
Site 4	MAE	0.2051	0.3518	0.5251	0.2617	0.5131	0.5347
	MSE	0.0423	0.1179	0.2741	0.0751	0.2015	0.3241
	MAPE (%)	3.1553	5.1618	6.9144	3.8117	5.9614	7.1874

(c) The forecasting accuracy of BA-RBFNN with the decomposition algorithms SSA and FEEMD is lower than that of BA-GRNN with the decomposition algorithms SSA and FEEMD, mostly.

Table 9 illustrates the following:

(a) In the one-step to three-step predictions, SSA-BA-RBFNN decreases the MAPE values from four sites based on SSA-CG-BA-RBFNN are 6.6862%, 3.9412% and 3.1007%;11.081%, 8.1497% and 6.4277%; 19.169%, 8.6555% and 10.541% and 21.384%, 9.3646% and 5.4413%, respectively.

(b) For the data from four sites, SSA-BA-GRNN decreases the MAPE values in the one-step to three-step predictions on the basis of SSA-CG-BA-GRNN are 15.431%, 2.7884% and 10.739%;28.472%, 7.0181% and 5.9328%; 20.456%, 20.182% and 10.733% and 24.763%, 14.197% and 9.5735%, respectively.

(c) The ξ_{MAPE} (%) values presents that from four sites in the one-step to three-step predictions, the MAPE values are decreased with 16.257%, 9.5431% and 3.3758%; 16.228%, 11.5025% and 5.8722%; 24.917%, 12.969% and 7.3186% and 24.724%, 12.641% and 7.3531% with FEEMD-CG-BA-RBFNN based on FEEMD-BA-RBFNN, respectively.

Table 12
Improvement percentages between four proposed hybrid models and RBFNN and GRNN.

		SSA-CG-BA-RBFNN *vs.* RBFNN			FEEMD-CG-BA- RBFNN *vs.* RBFNN			SSA-CG-BA-GRNN *vs.* GRNN			FEEMD-CG-BA-GRNN *vs.* GRNN		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	ξ_{MAE} (%)	−105.408	−65.2832	−42.1579	−166.619	−81.8014	−70.4799	−77.7459	−57.6779	−34.2896	−146.491	−77.4232	−54.1651
	ξ_{MSE} (%)	−323.076	−173.654	−102.181	−612.963	−231.959	−191.074	−215.573	−148.329	−80.2457	−511.111	−215.686	−137.973
	ξ_{MAPE} (%)	−105.591	−65.3882	−42.2669	−166.848	−81.9731	−70.6989	−78.1504	−57.9035	−34.4856	−146.822	−77.5652	−54.3086
Site 2	ξ_{MAE} (%)	−120.474	−57.541	−50.817	−188.357	−92.134	−89.569	−135.127	−67.737	−39.302	−158.585	−80.489	−64.523
	ξ_{MSE} (%)	−386.885	−148.108	−127.515	−736.619	−269.169	−258.537	−455.141	−181.595	−93.988	−567.416	−226.303	−170.829
	ξ_{MAPE} (%)	−120.131	−57.642	−50.887	−188.299	−92.008	−89.727	−135.052	−67.662	−39.291	−158.426	−80.597	−64.326
Site 3	ξ_{MAE} (%)	−158.422	−77.722	−49.738	−184.841	−122.279	−87.592	−153.822	−83.404	−39.775	−168.439	−86.830	−55.495
	ξ_{MSE} (%)	−565.714	−215.505	−123.961	−712.791	−393.461	−252.704	−547.222	−236.617	−95.401	−620.619	−248.941	−141.933
	ξ_{MAPE} (%)	−158.419	−77.701	−49.756	−184.582	−122.397	−87.416	−154.029	−83.256	−39.823	−168.618	−86.725	−55.488
Site 4	ξ_{MAE} (%)	−156.843	−86.001	−41.949	−179.761	−100.837	−81.202	−146.737	−81.208	−41.482	−163.692	−84.154	−77.644
	ξ_{MSE} (%)	−560.526	−245.621	−101.597	−684.375	−303.838	−228.416	−512.195	−228.249	−100.132	−597.222	−238.819	−215.746
	ξ_{MAPE} (%)	−156.722	−85.966	−41.897	−179.614	−101.027	−81.134	−146.459	−81.231	−41.517	−163.562	−84.141	−61.131

Table 13
Improvement percentages between four proposed hybrid models and ARIMA.

		SSA-CG-BA-RBFNN *vs.* ARIMA			FEEMD-CG-BA- RBFNN *vs.* ARIMA			SSA-CG-BA-GRNN *vs.* ARIMA			FEEMD-CG-BA-GRNN *vs.* ARIMA		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	ξ_{MAE} (%)	221.412	147.784	116.899	317.192	172.548	160.111	178.128	136.383	104.893	285.695	165.984	135.218
	ξ_{MSE} (%)	936.263	514.447	370.149	1646.29	645.361	576.859	672.951	457.583	319.142	1396.82	608.823	453.378
	ξ_{MAPE} (%)	221.677	147.985	117.011	317.523	172.853	160.381	178.743	136.762	105.141	286.191	166.243	135.379
Site 2	ξ_{MAE} (%)	239.367	143.201	127.011	343.855	196.604	185.341	261.919	158.942	109.677	298.027	178.628	147.641
	ξ_{MSE} (%)	1051.63	490.810	415.563	1878.87	779.089	712.466	1213.08	570.552	339.589	1478.65	677.014	513.715
	ξ_{MAPE} (%)	238.940	143.343	127.044	343.899	196.392	185.488	261.914	158.809	109.595	297.903	178.777	147.266
Site 3	ξ_{MAE} (%)	324.413	162.824	140.573	367.803	228.718	201.392	316.858	171.227	124.568	340.863	176.293	149.823
	ξ_{MSE} (%)	1696.19	590.069	477.715	2093.02	979.292	809.811	1646.29	636.245	404.041	1844.33	663.198	524.072
	ξ_{MAPE} (%)	323.933	162.826	140.647	366.852	228.935	201.163	316.731	171.043	124.685	340.663	176.174	149.858
Site 4	ξ_{MAE} (%)	304.495	167.122	130.746	340.588	188.429	194.554	288.579	160.239	129.986	315.282	164.469	188.771
	ξ_{MSE} (%)	1537.71	612.773	433.023	1844.79	732.836	768.329	1417.88	576.949	429.147	1628.71	598.747	734.827
	ξ_{MAPE} (%)	304.528	167.038	130.607	340.601	188.666	194.374	288.357	160.239	129.992	315.306	164.418	161.865

(d) In the one-step to three-step predictions, the MAPE values from four sites of by FEEMD-BA-GRNN are decreased 21.649%, 10.131% and 12.037%; 17.796%, 15.143% and 15.838%; 25.182%, 14.333% and 7.0716% and 25.229%, 9.4551% and 13.659% with FEEMD-CG-BA-GRNN, respectively.

Remark. By comparing SSA-CG-BA-RBFNN, SSA-CG-BA-GRNN, FEEMD-CG-BA-RBFNN and FEEMD-CG-BA-GRNN with SSA-BA-RBFNN, SSA-BA-GRNN, FEEMD-BA-RBFNN and FEEMD-BA-GRNN, the performance of the proposed optimization algorithm CG-BA is better than that of the original BA.

Tables 10–13 indicate the following:

(a) GRNN and ARIMA are the most and least accurate forecasting models, respectively, among RBFNN, GRNN and ARIMA in one-step to three-step prediction of the data from four sites.
(b) The four proposed models are more accurate than ELM and SVM. ELM is more accurate than RBFNN, GRNN, ARIMA and SVM.
(c) Based on the one-step to three-step prediction of RBFNN, from four sites, SSA-CG-BA-RBFNN decreases 105.591%, 65.3882% and 42.2669%; 120.131%, 57.6424% and 50.8865%; 158.419%, 77.7006% and 49.7562%; 156.722%, 85.9658% and 41.8967% MAPE values, respectively. And FEEMD-CG-BA-RBFNN decreases 166.848%, 81.9731% and 70.6989%; 188.299%, 92.0083% and 89.7268%; 184.582%, 122.397% and 87.4155%; 179.614%, 101.027% and 81.1342% MAPE values, respectively.
(d) In the one-step to three-step prediction, the MAPE promoted percentages from four sites of SSA-CG-BA-GRNN and FEEMD-CG-BA-GRNN by GRNN are −78.1504%, −57.9035% and −34.4856% and −146.822%, −77.5652% and −54.3086%; −135.052%, −67.6619% and −39.2909% and −158.426%, −80.5974% and −64.3257%; −154.029%, −83.2562% and −39.8232% and −168.618%, −86.7252% and −55.4882%; −146.459%, −81.2308% and −41.5171% and −163.562%, −84.1414% and −61.13065%, respectively.
(e) For Site 1, the MAPE promoted percentages, in the one-step to three-step prediction, from four sites of ARIMA with SSA-CG-BA-RBFNN, FEEMD-CG-BA-RBFNN, SSA-CG-BA-GRNN and FEEMD-CG-BA-GRNN are −1221.677%, −147.985% and −117.011%; −317.523%, −172.853% and −160.381%; −178.743%, −136.762% and −105.141%; −286.191%, −166.243% and −135.379%, respectively.
(f) For Site 2, the MAPE promoted percentages, in the one-step to three-step prediction, from four sites of ARIMA with SSA-CG-BA-RBFNN, FEEMD-CG-BA-RBFNN, SSA-CG-BA-GRNN

Table 14
Forecasting results of the persistence prediction test.

No.	Actual value (m/s)	Forecasting value (m/s)	MAPE (%)	No.	Actual value (m/s)	Forecasting value (m/s)	MAPE (%)
1	3.8	3.8418	1.0988	51	4.5	4.4465	1.1881
2	3.8	3.7554	1.1742	52	4.8	4.7456	1.1330
3	3.8	3.7536	1.2200	53	4.6	4.6508	1.1038
4	3.9	3.8576	1.0867	54	4.1	4.0504	1.2091
5	4.1	4.1469	1.1439	55	4.3	4.3519	1.2079
6	3.7	3.6552	1.2109	56	4.4	4.3468	1.2095
7	3.1	3.1386	1.2440	57	4.3	4.2474	1.2242
8	3.2	3.1592	1.2737	58	4.2	4.1463	1.2787
9	2.2	2.1748	1.1450	59	4.4	4.4485	1.1012
10	2.1	2.1252	1.2022	60	4.3	4.3470	1.0927
11	2.2	2.1744	1.1647	61	5.2	5.2608	1.1697
12	2	2.0227	1.1333	62	6.2	6.2764	1.2327
13	2.5	2.5284	1.1362	63	6.4	6.3210	1.2344
14	2.4	2.4285	1.1854	64	6.2	6.1210	1.2745
15	2.3	2.3289	1.2551	65	5.4	5.4598	1.1078
16	2.8	2.7645	1.2687	66	4.2	4.1539	1.0988
17	3.5	3.4555	1.2715	67	4.1	4.0514	1.1861
18	4	4.0486	1.2152	68	4.4	4.3482	1.1770
19	4.2	4.2510	1.2144	69	5.4	5.4656	1.2143
20	3.8	3.7584	1.0936	70	5.9	5.8301	1.1840
21	2.8	2.8315	1.1248	71	6	6.0666	1.1100
22	2.8	2.7650	1.2489	72	6	5.9321	1.1324
23	3.1	3.1383	1.2361	73	6.4	6.4788	1.2310
24	3.1	3.0665	1.0813	74	5.9	5.9689	1.1685
25	3.6	3.5583	1.1574	75	6.3	6.2274	1.1518
26	5.6	5.5395	1.0802	76	6.1	6.0293	1.1589
27	5	5.0582	1.1649	77	6	5.9268	1.2208
28	4.8	4.8592	1.2340	78	6.1	6.1661	1.0839
29	4.6	4.6569	1.2369	79	6.3	6.3734	1.1649
30	3.4	3.4370	1.0872	80	6.4	6.4716	1.1194
31	4.3	4.3526	1.2244	81	6.8	6.7207	1.1660
32	4.6	4.6511	1.1105	82	7.1	7.0178	1.1582
33	4.3	4.3517	1.2015	83	7.1	7.0177	1.1594
34	4.4	4.4540	1.2277	84	6.9	6.8151	1.2310
35	4.3	4.3543	1.2635	85	6.2	6.2696	1.1232
36	3.4	3.4419	1.2331	86	5.7	5.6276	1.2699
37	2.4	2.4273	1.1375	87	5.7	5.7692	1.2143
38	2.3	2.3275	1.1952	88	5.6	5.6698	1.2467
39	2.6	2.5691	1.1893	89	5.7	5.6365	1.1135
40	2.8	2.8338	1.2089	90	5.8	5.7259	1.2780
41	2.6	2.5684	1.2158	91	5.6	5.5296	1.2569
42	2.7	2.6657	1.2690	92	5.9	5.8345	1.1110
43	2.8	2.8342	1.2219	93	5.8	5.8674	1.1614
44	3.6	3.6397	1.1039	94	5.6	5.5303	1.2451
45	4.1	4.0520	1.1700	95	5.3	5.2394	1.1437
46	4.6	4.6558	1.2124	96	5.2	5.1429	1.0980
47	4.4	4.3494	1.1500	97	5.3	5.3587	1.1073
48	3.9	3.8546	1.1632	98	5.4	5.3363	1.1790
49	4.1	4.0489	1.2466	99	5.4	5.4637	1.1790
50	3.7	3.7445	1.2027	100	4.8	4.8524	1.0910

and FEEMD-CG-BA-GRNN are −238.940%, −143.343% and −127.044%; −343.899%, −196.392% and −185.488%; −261.914%, −158.809% and −109.595% and −297.903%, −178.777% and −147.266%, respectively.

(g) The MAPE promoted percentages, for Site 3 in the one-step to three-step prediction, from four sites of ARIMA with SSA-CG-BA-RBFNN, FEEMD-CG-BA-RBFNN, SSA-CG-BA-GRNN and FEEMD-CG-BA-GRNN are −323.933%, −162.826% and −140.647%; −366.852%, −228.935% and −201.163%; −316.731%, −171.043% and −124.685% and −340.663%, −176.174% and −149.858%, respectively.

(h) In the one-step to three-step prediction for Site 4, the MAPE promoted percentages from four sites of ARIMA with SSA-CG-BA-RBFNN, FEEMD-CG-BA-RBFNN, SSA-CG-BA-GRNN and FEEMD-CG-BA-GRNN are −304.528%, −167.038% and −130.607%; −340.601%, −188.666% and −194.374%; −288.357%, −160.239% and −129.992%; −315.306%, −164.418% and −161.865%, respectively.

Remark. By comparing SSA-CG-BA-RBFNN, SSA-CG-BA-GRNN, FEEMD-CG-BA-RBFNN and FEEMD-CG-BA-GRNN with ELM, SVM, RBFNN, GRNN and ARIMA, the forecasting performance of the proposed four hybrid models are better than that of the single models.

6.4. Experiment III: Persistence prediction test

To evaluate the proposed model, in this part a persistence prediction test is employed. In this test, the proposed model is used to output 100 continuous data. The forecasting results presented in Table 14 show that SSA-CG-BA-GRNN could always keep a high forecasting performance in this test.

6.5. Experiment IV: Diebold-Mariano (DM)-test and forecasting validity degree (FVD)

The DM test and FVD are conducted to further evaluate the levels of accuracy achieved by the proposed hybrid models (as shown in Fig. 7).

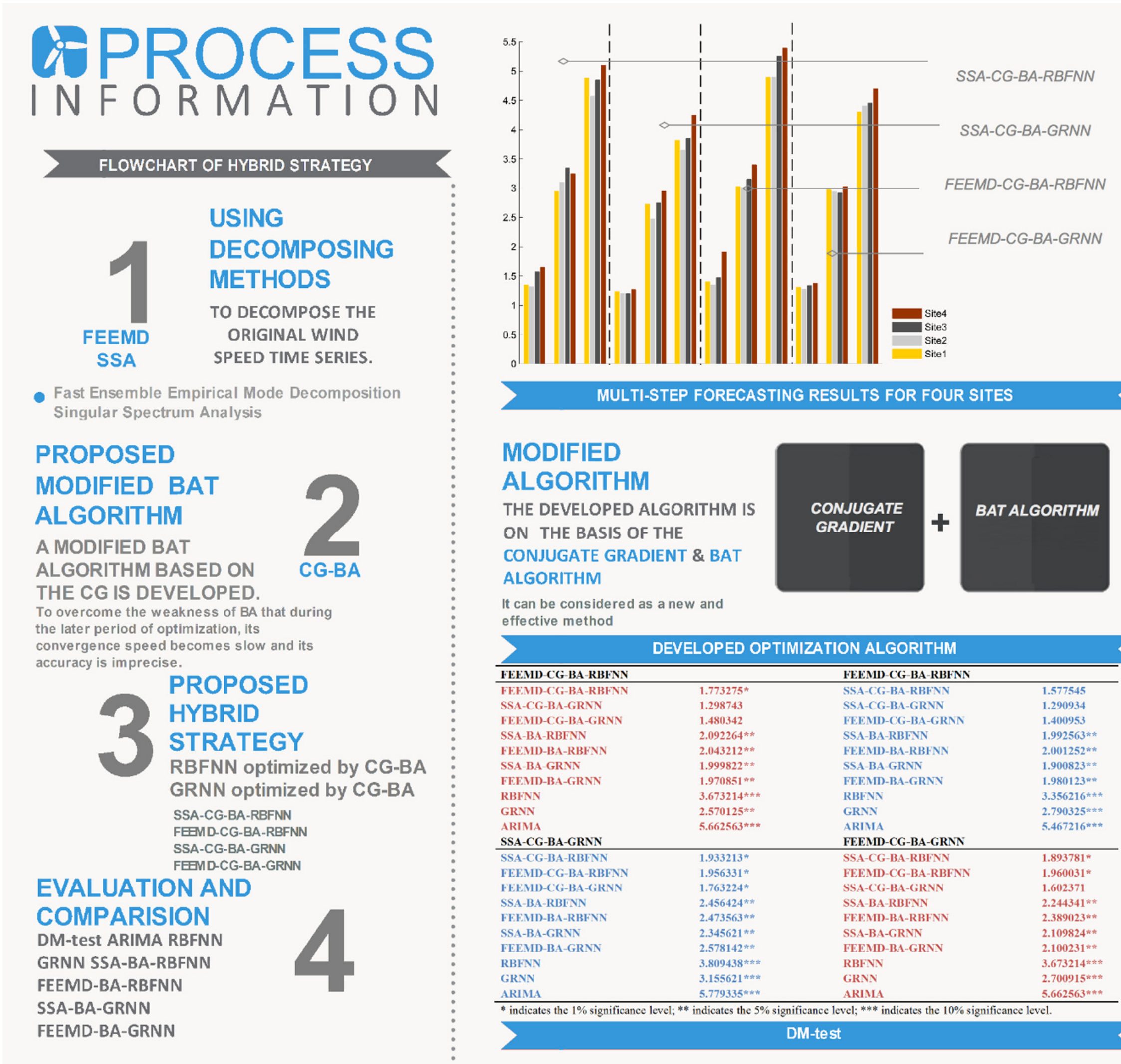


Fig. 7. Process of hybrid forecasting strategy and DM test results.

DM test which is a comparison test focusing on predictive accuracy, could be used to compare the forecasting performance of the proposed hybrid model with others. For more details, one can refer [55], which provides a complete description of the DM test theory.

FVD can be measured not only by the square sum of forecasting error but also by the mean and mean squared deviation of the forecasting accuracy. It is a useful tool to evaluate the forecasting accuracy of the model. For more details, one can refer [56], which provides a complete description of the DM test theory.

The results given in Table 15 indicate the following:

(a) SSA-CG-BA-RBFNN is more accurate than RBFNN and ARIMA at the 10% significance level, more accurate than SSA-BA-RBFNN, FEEMD-BA-RBFNN, SSA-BA-GRNN, FEEMD-BA-GRNN and GRNN at the 5% significance level, and more accurate than FEEMD-CG-BA-RBFNN at the 1% significance level.
(b) FEEMD-CG-BA-RBFNN is more accurate than RBFNN, GRNN and ARIMA at the 10% significance level and more accurate than SSA-BA-RBFNN, FEEMD-BA-RBFNN, SSA-BA-GRNN and FEEMD-BA-GRNN at the 5% significance level.
(c) SSA-CG-BA-GRNN is the most accurate model among these models. It is more accurate than SSA-CG-BA-RBFNN, FEEMD-CG-BA-RBFNN and FEEMD-CG-BA-GRNN at the 1% significance level, more accurate than SSA-BA-RBFNN, FEEMD-BA-RBFNN, SSA-BA-GRNN and FEEMD-CG-BA-GRNN at the 5% significance level, and more accurate than RBFNN, GRNN and ARIMA at the 10% significance level.
(d) FEEMD-CG-BA-GRNN is more accurate than SSA-CG-BA-RBFNN and FEEMD-CG-BA-RBFNN at the 1% significance level, more accurate than SSA-BA-RBFNN, FEEMD-BA-RBFNN, SSA-BA-GRNN and FEEMD-CG-BA-GRNN at the 5% significance level, and more accurate than RBFNN, GRNN and ARIMA at the 10% significance level.

FVD is measured to evaluate the forecasting accuracy of the hybrid models and the other six comparison models. A more accurate forecasting model leads to a larger FVD value. The results presented in Table 16 show that the FVD value for the SSA-CG-BA-GRNN model is larger than those of the comparison models.

Remark. Based on the results from the above two methods, the forecasting performance of SSA-CG-BA-GRNN has been globally evaluated. From the results of the DM test and FVD, one can see that SSA-CG-BA-GRNN is the most accurate forecasting model among the proposed forecasting architecture for multi-step wind speed forecasting. As shown in Table 17, SSA-CG-BA-GRNN is more

Table 15
Results for the DM test.

	SSA-CG-BA-RBFNN
DM-test	
FEEMD-CG-BA-RBFNN	1.773275*
SSA-CG-BA-GRNN	1.298743
FEEMD-CG-BA-GRNN	1.480342
SSA-BA-RBFNN	2.092264**
FEEMD-BA-RBFNN	2.043212**
SSA-BA-GRNN	1.999822**
FEEMD-BA-GRNN	1.970851**
RBFNN	3.673214***
GRNN	2.570125**
ARIMA	5.662563***
	FEEMD-CG-BA-RBFNN
SSA-CG-BA-RBFNN	1.577545
SSA-CG-BA-GRNN	1.290934
FEEMD-CG-BA-GRNN	1.400953
SSA-BA-RBFNN	1.992563**
FEEMD-BA-RBFNN	2.001252**
SSA-BA-GRNN	1.900823**
FEEMD-BA-GRNN	1.980123**
RBFNN	3.356216***
GRNN	2.790325***
ARIMA	5.467216***
	SSA-CG-BA-GRNN
SSA-CG-BA-RBFNN	1.933213*
FEEMD-CG-BA-RBFNN	1.956331*
FEEMD-CG-BA-GRNN	1.763224*
SSA-BA-RBFNN	2.456424**
FEEMD-BA-RBFNN	2.473563**
SSA-BA-GRNN	2.345621**
FEEMD-BA-GRNN	2.578142**
RBFNN	3.809438***
GRNN	3.155621***
ARIMA	5.779335***
	FEEMD-CG-BA-GRNN
SSA-CG-BA-RBFNN	1.893781*
FEEMD-CG-BA-RBFNN	1.960031*
SSA-CG-BA-GRNN	1.602371
SSA-BA-RBFNN	2.244341**
FEEMD-BA-RBFNN	2.389023**
SSA-BA-GRNN	2.109824**
FEEMD-BA-GRNN	2.100231**
RBFNN	3.673214***
GRNN	2.700915***
ARIMA	5.662563***

* 1% significance level.
** 5% significance level.
*** 10% significance level.

Table 16
Results for FVD.

Model	FVD		
	1-step	2-step	3-step
SSA-CG-BA-RBFNN	98.5264	96.8396	95.1481
FEEMD-CG-BA-RBFNN	98.4681	96.8553	94.8889
SSA-CG-BA-GRNN	**98.7722**	**97.2742**	**96.1030**
FEEMD-CG-BA-GRNN	98.6750	97.0318	95.5347
SSA-BA-RBFNN	98.2778	96.5825	94.8191
FEEMD-BA-RBFNN	98.0774	96.4420	94.5660
SSA-BA-GRNN	98.4156	96.9320	95.7032
FEEMD-BA-GRNN	98.2897	96.6151	94.9104
RBFNN	95.4801	94.0499	92.1663
GRNN	96.5663	94.5909	92.9118
ARIMA	94.5746	91.9461	88.9130

The best results are formatted in bold.

Table 17
Total computation time of each model.

Model	CPU time (s)
SSA-CG-BA-RBFNN	38.6145
FEEMD-CG-BA-RBFNN	39.1547
FEEMD-CG-BA-GRNN	31.0245
SSA-CG-BA-GRNN	28.6414
SSA-BA-RBFNN	23.6554
FEEMD-BA-RBFNN	24.0046
SSA-BA-GRNN	22.0894
FEEMD-BA-GRNN	23.1663
RBFNN	16.2461
GRNN	14.2541
ARIMA	8.9564
ELM	12.6791
SVM	18.4989

efficient than FEEMD-CG-BA-GRNN, SSA-CG-BA-RBFNN and FEEMD-CG-BA-RBFNN. The relation between the wind speed and wind power generation could be expressed as the following equation:

$$P_a = \left\{ \frac{\exp[-(v_c/c)^k] - \exp[-(v_r/c)^k]}{(v_r/c)^k - (v_c/c)^k} - \exp[-(v_f/c)^k] \right\} \times P_r \qquad (2)$$

where P_a is the average power output of the wind turbine (kW), P_r is the rated electrical power of the wind turbine (kW), v_c is the cut-in wind speed (m/s), v_f is the cut-off wind speed (m/s), v_r is the nominal wind speed (m/s), and c is the Weibull scale parameter (m/s). It is observed that the accurate wind speed forecasting plays an important role in the wind power generation.

7. Conclusion

As one of the most promising potential renewable energies, wind energy has been a focus of many scientists and researchers and supported by almost every government across the world. To integrate wind energy into the power system, it is important to forecast wind power generation. Wind speed is affected by various environmental factors, so wind speed data present high fluctuations, autocorrelation and stochastic volatility, and it is difficult to forecast wind speed using a single model. In this paper, four hybrid models based on two decomposition algorithms, SSA and FEEMD, and two neural networks, RBFNN and GRNN, are proposed for multi-step wind speed forecasting. Meanwhile, to improve the performance of the neural networks, a new improved BA algorithm, CG-BA, based on CG is proposed to optimize the initial weights and thresholds of neural networks. Based on a series of forecasting results, the DM test and FVD, the following can be concluded: (a) the hybrid SSA-CG-BA-GRNN model is the most accurate model among the four proposed models in multi-step wind speed forecasting; (b) the decomposition algorithm SSA is better than FEEMD in this study; (c) the performance of the single model GRNN is more accurate than RBFNN and ARIMA.

Thus, the proposed SSA-CG-BA-GRNN model, which has the highest precision, is a promising model for use in the future. This hybrid model can also be applied in many other fields, such as tourism demand forecasting, product sales forecasting, power load forecasting, and traffic flow forecasting.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61331007 and 61471105) and 973 Project (No. 613273).

Appendix A

A.1. SSA and FEEMD

SSA is a novel nonparametric method, which is employed in the analysis of time series and combines multivariate statistic and probability theory, and it is often used for identifying and extracting periodic, quasi-periodic and oscillatory components from the primal data [28]. Standard SSA performs four steps, which include embedding, singular value decomposition, grouping and diagonal averaging. However, the first two steps are also called the time series decomposition, and steps three and four are known as the reconstruction. For more details, one can refer [28,44], which provides a complete description of the SSA theory.

FEEMD is an extension of the empirical mode decomposition [45] and ensemble empirical mode decomposition techniques [46]. The fast ensemble empirical mode decomposition technique is a time-domain decomposing method, which can convert a group of time series into multiple empirical modes, named as the intrinsic mode functions (IMFs). $y(t)$ is the time series, it can be decomposed and expressed using the following formula:

$$y(t) = \sum_{j=1}^{n} \mathbf{IMF}_j(t) + r_n(t) \tag{A1}$$

where $\mathbf{IMF}_j\,(t), j = 1, 2, \ldots, n$, is the intrinsic mode function (i.e., local oscillation) based on empirical mode decomposition and $r_n(t)$ is the nth residue (i.e., local trend). For more details, one can refer [47], which provides a complete description of the SSA theory.

Appendix B

B.1. Standard RBFNN

Definition 1. Cluster centers were composed of elements $m_i^j\ (j = 1 - n)$ from the center vector m_i, which is in the input space.

Definition 2. With elements $I_j\ (j = 1 - n)$, the distance measure is used to determine how far the center vector m^i is from an input vector I. The popular distance measure is the Euclidian distance, defined as

$$d_i = \sqrt{\sum\nolimits_{j=1}^{n} k_j^i (I_j - m_j^i)^2} \tag{B1}$$

where k_j^i is the (i, j)th element of the shape matrix K, defined as the inverse of the covariance matrix:

$$k_j^i = \frac{h_j^i}{(\sigma_j^i)^2} \tag{B2}$$

where h_j^i is the correlation coefficient, and σ_j^i represents the marginal standard deviation.

Definition 3. A transfer function transforms the Euclidian summation $d_j\ (i = 1 - m)$ and gives an output for each node. The output generated by the hidden layer was from the input layer via a distance measure of Eq. (B1) and a transfer function. A weighted sum of the outputs of $\phi(d_i)$ from the hidden layer processed the output of the network, i.e.,

$$O = w_0 + \sum_{i=1}^{m} w_i \phi(d_i) \tag{B3}$$

Appendix B2 Standard GRNN.

Definition 1. The input layer accepts information and also stores an input vector X, whose dimension m equals the number of input layer neurons. The pattern layer is then fed the data that comes from the input neurons of the input layer. A nonlinear transformation, which transforms the input space into the pattern space, was used by the pattern layer. The neurons of the pattern layer can remember the relation between the input neuron and the proper response of the pattern layer, in which the number of neurons is equal to the number of training samples n. The pattern Gaussian function of p_i is expressed as

$$p_i = \exp\left[-\frac{(X - X_i)^T (X - X_i)}{2\sigma^2}\right] \quad (i = 1, 2, \ldots, n) \tag{B4}$$

where σ represents the spread parameter, and X is the network's input variable. In the pattern layer, X_i is a specific training sample of neuron i.

Definition 2. S_s and S_w are the two summations of the summation layer. The simple summation S_s is used to calculate the arithmetic sum from the outputs that belong to the pattern layer, and i is the interconnection weight of the simple summation. The weighted summation S_w is used to calculate the weighted sum from the outputs that belong to the pattern layer, and w is the interconnection weight of the weighted summation. The transfer functions can be described by Eqs. (B4) and (B5):

$$S_s = \sum_{t=1} p_t, \quad t = 1 \ldots n \tag{B5}$$

$$S_{wt} = \sum_{t=1} w_t p_t, \quad t = 1 \ldots n \tag{B6}$$

where w_t is the weight of pattern neuron t that is connected to the summation layer.

Definition 3. In the output layer, the number of neurons is equal to the dimension k of the output vector Y. In the summation layer, after the summations, the output absorbs the neurons, and the output Y of the output neurons can be computed as

$$\hat{Y}_o = S_s / S_{wo}, \quad o = 1 \ldots k \tag{B7}$$

If the training set is given, the spread parameter R is the only parameter that must be confirmed.

Appendix C

C.1. Accuracy estimating indexes

The detailed equations of MAE MSE and MAPE are given in Table C1.

Table C1
Three metric rules.

Metric	Definition	Equation
MAE	The average absolute forecast error of n times forecast results	$MAE = \frac{1}{N}\sum_{n=1}^{N} \lvert y_n - \hat{y}_n \rvert$
MSE	The average of the prediction error squares	$MSE = \frac{1}{N}\sum_{n=1}^{N} (y_n - \hat{y}_n)^2$
MAPE	The average of absolute error	$MAPE = \frac{1}{N}\sum_{n=1}^{N} \left\lvert \frac{y_n - \hat{y}_n}{y_n} \right\rvert \times 100\%$

Where y_n and $\hat{y}_n$ denote the actual value and predicted value, respectively, of the nth data for the performance estimate, and N is the length of the dataset to compare and evaluate.

Additionally, to obtain the detailed promoting percentages when comparing two forecasting, i.e. model 1 and model 2, three percentage error indexes are also defined as follows:

$$\xi_{MAE} = \frac{MAE_2 - MAE_1}{MAE_1} \times 100\% \tag{C1}$$

$$\xi_{MSE} = \frac{MSE_2 - MSE_1}{MSE_1} \times 100\% \tag{C2}$$

$$\xi_{MAPE} = \frac{MAPE_2 - MAPE_1}{MAPE_1} \times 100\% \tag{C3}$$

The negative value of ξ_{MAE} (%) means model 2 decreases $|\xi_{MAE}|$% MAE value based on model 1, the positive value of ξ_{MAE} (%) means model 2 increases $|\xi_{MAE}|$% MAE value based on model 1. So do ξ_{MSE} (%) and ξ_{MAPE} (%).

References

[1] Yesilbudak M, Sagiroglu S, Colak I. A new approach to very short term wind speed prediction using k-nearest neighbor classification. Energy Convers Manage 2013;69(69):77–86.
[2] Meng A, Ge J, Yin H, Chen S. Wind speed forecasting based on wavelet packet decomposition and artificial neural networks trained by crisscross optimization algorithm. Energy Convers Manage 2016;114:75–88.
[3] Ozay C, Celiktas MS. Statistical analysis of wind speed using two-parameter Weibull distribution in Alaçatı region. Energy Convers Manage 2016;121:49–54.
[4] Liu H, Tian HQ, Li YF, Zhang L. Comparison of four Adaboost algorithm based artificial neural networks in wind speed predictions. Energy Convers Manage 2015;92(92):67–81.
[5] Zhao ZY, Hu J, Zuo J. Performance of wind power industry development in China: a DiamondModel study. Renew Energy 2009;34(12):2883–91.
[6] Zhang C, Wei H, Zhao X, Liu T, Zhang K. A Gaussian process regression based hybrid approach for short-term wind speed prediction. Energy Convers Manage 2016;126:1084–92.
[7] Liu H, Tian H, Li Y. Four wind speed multi-step forecasting models using extreme learning machines and signal decomposing algorithms. Energy Convers Manage 2015;100:16–22.
[8] Zhou J, Jing S, Gong L. Fine tuning support vector machines for short-term wind speed forecasting. Energy Convers Manage 2011;52(4):1990–8.
[9] Lei M, Shiyan L, Chuanwen J, Hongling L, Yan Z. A review on the forecasting of wind speed and generated power. Renew Sustain Energy Rev 2009;13(4):915–20.
[10] Landberg L. Short-term prediction of the power production from wind farms. J Wind Eng Ind Aerodyn 1999;80:207–20.
[11] Alexiadis MC, Dokopoulos PS, Sahsamanoglou HS, Manousaridis IM. Short term forecasting of wind speed and related electrical power. Sol Energy 1998;63(1):61–8.
[12] Negnevitsky M, Potter CW. Innovative short-term wind generation prediction techniques. In: Proceedings of the power systems conference and exposition. p. 60–5.
[13] Riahy GH, Abedi M. Short term wind speed forecasting for wind turbine applications using linear prediction method. Renew Energy 2008;33(1):35–41.
[14] Ma L, Luan SY, Jiang CW, Liu HL, Zhang Y. A review on the forecasting of wind speed and generated power. Renew Sustain Energy Rev 2009;13:915–20.
[15] Xiao L, Wang J, Yang X, Xiao L. A hybrid model based on data preprocessing for electrical power forecasting. Int J Electr Power Energy Syst 2015;64(64):311–27.
[16] Lydia M, Kumar SS, Selvakumar AI, Kumar GEP. Linear and non-linear autoregressive models for short-term wind speed forecasting. Energy Convers Manage 2016;112:115–24.
[17] Noorollahi Y, Jokar MA, Kalhor A. Using artificial neural networks for temporal and spatial wind speed forecasting in Iran. Energy Convers Manage 2016;115:17–25.
[18] Barbounis TG, Theocharis JB. A locally recurrent fuzzy neural network with application to the wind speed prediction using spatial correlation. Neurocomputing 2007;70(7/9):1525–42.
[19] Focken U, Lange M, Moonnich K, Waldl H-P, Georg Beyer H, Luig A. Short-term prediction of the aggregated power output of wind farms – a statistical analysis of the reduction of the prediction error by spatial smoothing effects. J Wind Eng Ind Aerodyn 2002;90(3):231–46.
[20] Niu M, Sun S, Wu J, Yu L, Wang J. An innovative integrated model using the singular spectrum analysis and nonlinear multi-layer perceptron network optimized by hybrid intelligent algorithm for short-term load forecasting. Appl Math Model 2016;40(5–6):4079–93.
[21] Xiao L, Shao W, Wang C, Zhang K, Lu H. Research and application of a hybrid model based on multi-objective optimization for electrical load forecasting. Appl Energy 2016;180(C):213–33.
[22] Su Z, Wang J, Lu H, Zhao G. A new hybrid model optimized by an intelligent optimization algorithm for wind speed forecasting. Energy Convers Manage 2014;85(9):443–52.
[23] Monfared M, Rastegar H, Kojabadi HM. A new strategy for wind speed forecasting using artificial intelligent methods. Renew Energy 2009;34:845–8.
[24] Cadenas E, Rivera W. Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks. Renew Energy 2009;34(1):274–8.
[25] Xiao L, Shao W, Liang T, Wang C. A combined model based on multiple seasonal patterns and modified firefly algorithm for electrical load forecasting. Appl Energy 2016;167:135–53.
[26] Sfetsos A. A comparison of various forecasting techniques applied to mean hourly wind speed time series. Renew Energy 2000;21(1):23–35.
[27] Mohandes MA, Halawani TO, Rehman S, Hussain AA. Support vector machines for wind speed prediction. Renew Energy 2004;29:939–47.
[28] Chen N, Qian Z, Meng X. Multi-step wind speed forecasting based on wavelet and gaussian processes. Math Probl Eng 2013.
[29] Liu H, Tian H-Q, Pan D-F, Li Y-F. Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks. Appl Energy 2013;107:191–208.
[30] Liu H, Chen C, Tian H-Q, Li Y-F. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. Renew Energy 2012;48:545–56.
[31] Wang Y, Wang S, Zhang N. A novel wind speed forecasting method based on ensemble empirical mode decomposition and GA-BP neural network. Presented at the IEEE power and energy society general meeting; 2013.
[32] Liu H, Tian H-Q, Li Y-F. Comparison of new hybrid FEEMD-MLP, FEEMD-ANFIS, Wavelet Packet-MLP and Wavelet Packet-ANFIS for wind speed predictions. Energy Convers Manage 2014;89:1–11.
[33] Liu H, Tian H-Q, Chen C, Li Y-F. An experimental investigation of two WaveletMLP hybrid frameworks for wind speed prediction using GA and PSO optimization. Int J Electr Power Energy Syst 2013;52(1):161–73.
[34] Poitras G, Cormier G. Wind speed prediction for a target station using neural networks and particle swarm optimization. Wind Eng 2011;35(3):369–80.
[35] Carro-Calvo L, Salcedo-Sanz S, Prieto L, Kirchner-Bossi N, Portilla-Figueras A, Jiménez-Fernández S. Wind speed reconstruction from synoptic pressure patterns using an evolutionary algorithm. Appl Energy 2012;89(1):347–54.
[36] Xiao L, Wang J, Hou R, Wu J. A combined model based on data pre-analysis and weight coefficients optimization for electrical load forecasting. Energy 2015;82:524–49.
[37] Yang XS. A new metaheuristic bat-inspired algorithm. Nature inspired cooperative strategies for optimization (NICSO 2010). Studies in computational intelligence, vol. 284. Berlin: Springer Verlag; 2010. p. 65–74.
[38] Yang XS. Bat algorithm: literature review and applications. Int J Bio-Inspired Comput 2013;5(3):141–9.
[39] Du ZY, Liu B. Image matching using a bat algorithm with mutation. Appl Mech Mater 2012;203:88–93.
[40] Mishra S, Shaw K, Mishra D. A new meta-heuristic bat inspired classification approach for microarray data. Procedia Technol 2012;4:802–6.
[41] Musikapun P, Pongcharoen P. Solving multi-stage multi-machine multiproduct scheduling problem using bat algorithm. Second international conference on management and artificial intelligence (IPEDR), vol. 35. Singapore: IACSIT Press; 2012. p. 98–102.
[42] Niknam T, Sharifinia S, Abaraghooee RA. A new enhanced bat-inspired algorithm for finding linear supply function equilibrium of GENCOs in the competitive electricity market. Energy Convers Manage 2013;76:1015–28.
[43] Sambariya DK, Prasad R. Robust tuning of power system stabilizer for small signal stability enhancement using metaheuristic bat algorithm. Int J Electr Power Energy Syst 2014;61:229–38.
[44] Skittides C, Früh WG. Wind forecasting using Principal Component Analysis. Renew Energy 2014;69(September):365–74.
[45] Huang NE, Liu HH. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. Proc Roy Soc A Math Phys Eng Sci 1971;454:903.
[46] Wu Z, Huang NE. Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 2009;1(01):1–41.
[47] Wang YH, Yeh CH, Young HWV, Hu K, Lo MT. On the computational complexity of the empirical mode decomposition algorithm. Physica A 2014;400(2):159–67.
[48] Yang XS. A new metaheuristic bat-inspired algorithm. Nature inspired cooperative strategies for optimization (NICSO), studies in computational intelligence, vol. 284. Springer; 2010. p. 65–74.
[49] Yang XS. Bat algorithm for multiobjective optimization. Int J Bio-Inspired Comput 2011;3(5):267–74.
[50] Shen F, Chao J, Zhao J. Forecasting exchange rate using deep belief networks and conjugate gradient method. Neurocomputing 2015;167:243–53.
[51] Guo XT, Zhu Y. Evolutionary neural networks based on genetic algorithms. J Tsinghua Univ (Sci Technol) 2000;40(10):116–9.
[52] Specht DF. A general regression neural network. IEEE Trans Neural Networks 1991;2(6):568–76.
[53] Polat O, Yildirim T. Hand geometry identification without feature extraction by general regression neural network. Expert Syst Appl 2008;34:845–9.
[54] Hecht-Nielsen SR. Kolmogorov's mapping neural network existence theorem. In: IEEE joint conf on neural networks, New York, USA, vol. 3; 1987. p. 11–4.
[55] Diebold FX, Mariano R. Comparing predictive accuracy. J Bus Econ Stat 1995;13:253–63.
[56] Chen HY, Hou DP. Research on superior combination forecasting model based on forecasting effective measure. J Univ Sci Technol China 2002;2:006.