Short-Term Wind Power Prediction Based on Hybrid Neural Network and Chaotic Shark Smell Optimization

2 authors:

Oveis Abedinia
Budapest University of Technology and Economics
86 PUBLICATIONS **2,333** CITATIONS

SEE PROFILE

Nima Amjady
Semnan University
252 PUBLICATIONS **10,061** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   yazdaninejad View project

Project   Innovative Approaches to Advanced Optimization in Electricity Markets View project

Short-Term Wind Power Prediction Based on Hybrid Neural Network and Chaotic Shark Smell Optimization

Oveis Abedinia[1,#] and Nima Amjady[1]

1 Department of Electrical Engineering, Semnan University, Semnan, Iran
Corresponding Author / E-mail: oveis.abedinia@semnan.ac.ir, TEL: +98-45-33722255, FAX: +98-45-33522255

By the quick growth of wind power generation in the world, this clean energy becomes an important green electrical source in many countries. However, volatile and non-dispatchable nature of this energy source motivates researchers to find accurate and robust methods to predict its future values. Because of nonlinear and complex behaviors of this signal, more efficient wind power forecast methods are still demanded. In this paper, a new forecasting engine based on Neural Network (NN) and a novel Chaotic Shark Smell Optimization (CSSO) algorithm is proposed. Choosing optimal number of nodes for the hidden layer can enhance the efficiency of the NN's training performance. Accordingly, a new meta-heuristic algorithm is presented in this paper, which is based on shark abilities in nature, for optimizing the number of hidden nodes pertaining to the NN. Effectiveness of the proposed forecasting strategy is tested on two real-world case studies for predicting wind power. The obtained results demonstrate the capability of the proposed technique to cope with the variability and intermittency of wind power time series for providing accurate predictions of its future values.

1. Introduction

With the quick growth of wind energy, the large-scale integration of this volatile energy source negatively impacts operation of electric power systems. Forecasting the wind power is one of the most important technologies against the challenges that large scale wind power integration brings to power system. Thus, more efficient wind power forecast methods are still demanded to further improve not only the forecast accuracy but also the risk assessment.[1,2] For this purpose, several models have been proposed by researchers to tackle wind power forecasting problem. In Ref. 3, the potential of Evolutionary Product Unit Neural Networks (EPUNNs) has been proposed to solve wind speed forecasting problem for wind farms. This reference has tested the proposed method on a Spanish test case. In Ref. 4, NN-based forecast engine has been presented for weather forecasting by considering the wind factor. In Ref. 5, combination of differential Empirical Mode Decomposition (EMD) and Relevance Vector Machine (RVM) has been presented for prediction of short-term wind power output using raw data of wind farm. In Ref. 6, two Neural Network (NN) based models are proposed for direct and rapid construction of prediction intervals for short-term wind power. Combination of nonparametric and time-varying regression and time-

series model, i.e. Holt-Winters and ARMA, applied for considering residual autocorrelation and seasonal dynamics, is proposed in Ref. 7 for prediction of wind power. A Hybrid Iterative Forecast Method (HIFM) for wind power forecasting is presented in Ref. 8. A two stage feature selection technique is also introduced for selecting the most relevant and the less redundant input variables in this reference. In Ref. 9, Ridgelet Neural Network (RNN), using Ridgelets as the activation functions in the hidden nodes, has been proposed as forecast engine to predict the wind power. Also, a new version of Differential Evolution (DE) algorithm with novel crossover operator and selection mechanism is presented in this reference to train the RNN. A review of wind power forecast methods can be found in Refs. 10-13.

In this paper, we propose a new structure for forecasting engine based on NN as well as a new stochastic search algorithm. It is clear that the random selection of hidden neurons may cause overfitting or underfitting problem in the NN. Thus, we also propose a new meta-heuristic algorithm, inspired from shark smell abilities as a superior hunter in nature, to determine the optimum number of neurons in the hidden layer. By applying this new algorithm, called Chaotic Shark Smell Optimization (CSSO), we optimize the parameters of NN-based forecasting engine to increase its training efficiency and forecast accuracy.
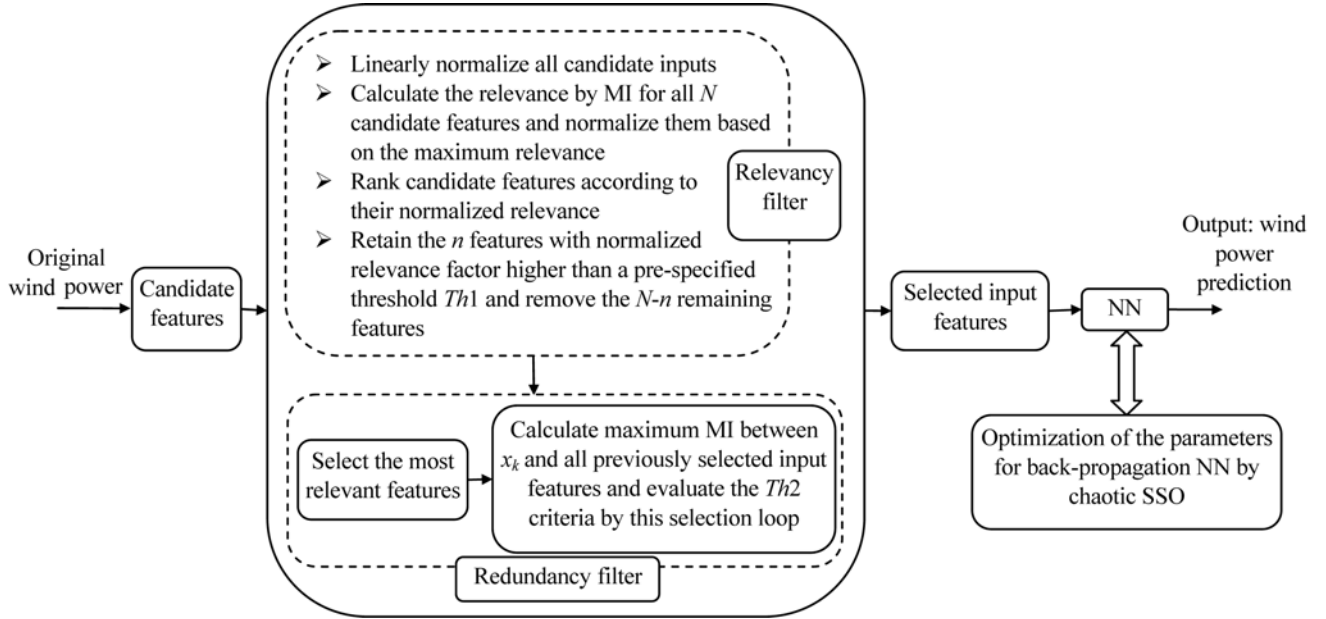
Springer

Fig. 1 The flowchart of the proposed wind power forecast strategy

The new contributions of this paper can be summarized as follows:

◉A new prediction strategy for short-term wind power forecasting is presented based on a new NN-based forecasting engine. In this model the parameters of NN, such as number of neurons in the hidden layer, are fine-tuned by an intelligent algorithm.

◉ A new meta-heuristic algorithm is proposed, which is based on shark smell optimization. In this paper, the chaotic version of this algorithm is presented, which improves the search abilities of the basic method in the search environment.

The rest of the paper is organized as follows. In Section 2, the structure of the proposed forecasting engine is introduced. Section 3 presents the suggested CSSO algorithm. Numerical results obtained from the proposed wind power forecasting strategy are presented and discussed in Section 4. Finally, Section 5 concludes the paper.

2. The Proposed Model for Wind Power Prediction

In this paper, back-propagation neural network is proposed as the forecasting engine. However, before describing the classic and modified versions of this forecasting engine, its inputs should be determined. The proposed NN is fed by the inputs selected by the double-filter feature selection method of Ref. 14, which is based on the information theoretic criteria of Mutual Information (MI) and Information Gain (IG). This method has two cascaded filters to filter out irrelevant candidate features (i.e. the inputs that have low mutual information with the output variable) and redundant candidate features (i.e. the inputs that have high mutual information with the other inputs), respectively. Only the relevant non-redundant candidate inputs, constituting a minimum subset of the most informative features for predicting the output variable, are selected by the feature selection method. As this feature selection method is not the focus of this paper, it is not further discussed here. The interested reader can refer to Ref. 14 for details of this technique. The main structure of proposed wind power forecast strategy is presented in Fig. 1.
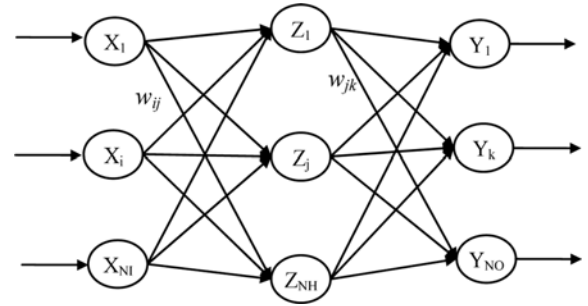

Fig. 2 Typical three-layer back propagation neural network

2.1 Back Propagation Neural Network

Neural networks have been applied successfully for different engineering problems in recent years. In this paper, we apply the back-propagation NN, which is one of the most popular NNs. The back-propagation learning algorithm includes four main steps as: Feed forward computation, Back propagation to the output layer, Back propagation to the hidden layer and weight updates. This forecasting engine has good abilities for dealing with nonlinear systems, such as forecasting problem of wind power.[15] In this problem, the next several hours' prediction is the main function of the forecast engine. The proposed optimization algorithm will be employed for improving this forecasting engine. The structure of a three layered back propagation NN is presented in Fig. 2. For each hidden neuron j, the input Ij and output Oj are defined as:

$$I_i = \sum_i w_{ij} O_i \qquad (1)$$

$$O_j = f(I_j + h_j) \qquad (2)$$

where w_{ij} is the weight between the ith neuron in the input layer and j^{th} neuron in the hidden layer; $f(.)$ is the activation function of the hidden neurons; O_i and O_j are the output of input neuron i and hidden neuron j; hj is the bias of hidden neuron j.

The initial number of neurons in the hidden layer is considered 7.

Moreover, in this model, the activation function is bipolar sigmoid and Purelin for the hidden and output layer, respectively. Also, Trainlm is selected as the training function.[16] The bipolar sigmoid activation function is presented as Ref. 16:

$$f(x) = \frac{2}{(1+\exp(-x))} - 1 \tag{3}$$

Also, the sum of absolute errors E and mean absolute error E_m in the output layer are evaluated as below:

$$E_m = \frac{1}{2n} \sum_{k=1}^{n} |T_k - O_k| \tag{4}$$

In the above equations, n is the number of training patterns; T_k and O_k are the target value and obtained value of kth neuron in the output layer. The back propagation error for k^{th} output node and j^{th} hidden node, denoted by δ_k and δ_j, respectively, can be calculated as given in Ref. 16. Using the back-propagated errors, the weights of the NN are updated to reduce the errors.[17,18]

2.2 Improved Back-Propagation NN

To improve the performance of the forecasting engine, the proposed CSSO is applied, which can be described as the following steps:

◉ Step 1: At first, the proposed CSSO algorithm is initialized. In this stage, the position and velocity of shark are determined. Decision variables of the CSSO algorithm, i.e. positions of shark, are selected as the weights and biases of the NN-based forecasting engine.

◉ Step 2: In the second step, the objective function of the proposed algorithm is evaluated. In this paper, mean absolute error E_m, defined in Eq. (5) is considered as the objective function of the NN in the training procedure. Here, T_k and O_k are actual and predicted wind powers, respectively.

◉ Step 3: Compare the value of the objective function with its value in the previous iteration of CSSO.

◉ Step 4: Evolve the position of shark based on the result of the comparison in the previous step. The evolution process of shark will be introduced in the next section.

◉ Step 5: If the stopping criteria is satisfied, go to the next step. Otherwise, go back to Step 2.

◉ Step 6: The individual of CSSO algorithm leading to the lowest objective function value (i.e. the lowest value of E_m) is selected as the CSSO solution. The weights and biases of this optimal solution are loaded to NN-based forecasting engine, which are used as the initial weight and bias values of the NN for applying the back-propagation learning algorithm. In other words, the back-propagation learning algorithm begins from these initial weight and bias values.

3. Chaotic Shark Smell Optimization

3.1 Shark Smell Optimization

Shark Smell Optimization (SSO) algorithm has recently been proposed by Abedinia et al. in Ref. 19. This algorithm is based on distinct shark smell abilities for localizing the prey. In sharks' movement, the concentration of the odor is an important factor to guide the shark to the prey. In other words, the shark moves in the way with higher odor concentration. Fig. 3 presents the movement of shark to the odor source based on its concentration. This characteristic is used in the proposed SSO algorithm to find the solution of an optimization problem. In this algorithm, we consider some assumptions, which have been presented in Ref. 19.

3.1.1 Initialization

The proposed SSO algorithm starts the search process when the shark smells an odor particle. Afterward, a population of the initial solutions for the optimization problem is randomly generated within the feasible search domain. Each solution represents one odor particle, i.e. one possible position for the shark. The initial population of SSO at the beginning of the search process can be represented as:

$$[X_1^1, X_2^1, ? \ X_{NP}^1], \ NP = \text{Popultaion size} \tag{5}$$

where the ith initial position vector X_i^1, i.e. i^{th} initial candidate solution for the optimization problem, is as follows:

$$X_i^1 = [x_{i,1}^1, x_{i,2}^1, ? \ x_{i,ND}^1], \ i = 1, \dots NP \tag{6}$$

where $x_{i,j}^1$ is the jth dimension of the ith shark position or equivalently j^{th} decision variable of the ith individual X_i^1; ND is the number of decision variables of the optimization problem. As described in the previous section, the decision variables of the optimization problem pertaining to NN training phase are the NN weights and biases. Magnitude of odor in each position indicates its closeness to the prey.

3.1.2 Evolution Process

Shark has an initial velocity and location in the starting point. By increasing the concentration of the Odor Particles (OPs), the position and velocity of shark will be changed. So, corresponding to the position vectors, we have NP initial velocity vectors as follows:

$$[V_1^1, V_2^1, ? \ V_{i,NP}^1] \tag{7}$$

where each velocity vector has components in all dimensions:

$$V_i^1 = [v_{i,1}^1, v_{i,2}^1, ? \ v_{i,ND}^1], \ i = 1, \dots, NP \tag{8}$$

By increasing the concentration, the velocity of shark will increase. The mathematical model for this type of movement can be expressed through gradient of the objective function, illustrating the direction that the objective function increases by the highest rate:

$$V_i^m = \mu_m.R1.\nabla(OF)\big|_{X_i^m}, \ i = 1, \dots, NP, \ m = 1, \dots, M \tag{9}$$

where OF represents the objective function and $\nabla(OF)$ is its gradient; μ_m indicates the gradient constant. In Eq. (9), m represents the stage number and M is the maximum number of stages for the forward movement of shark. Also, $\mu_m \in (0,1)$ and $R1$ is a random number with uniform distribution in the interval $(0,1)$. According to the Eq. (9), we can write the velocity of shark in every dimension as follows:
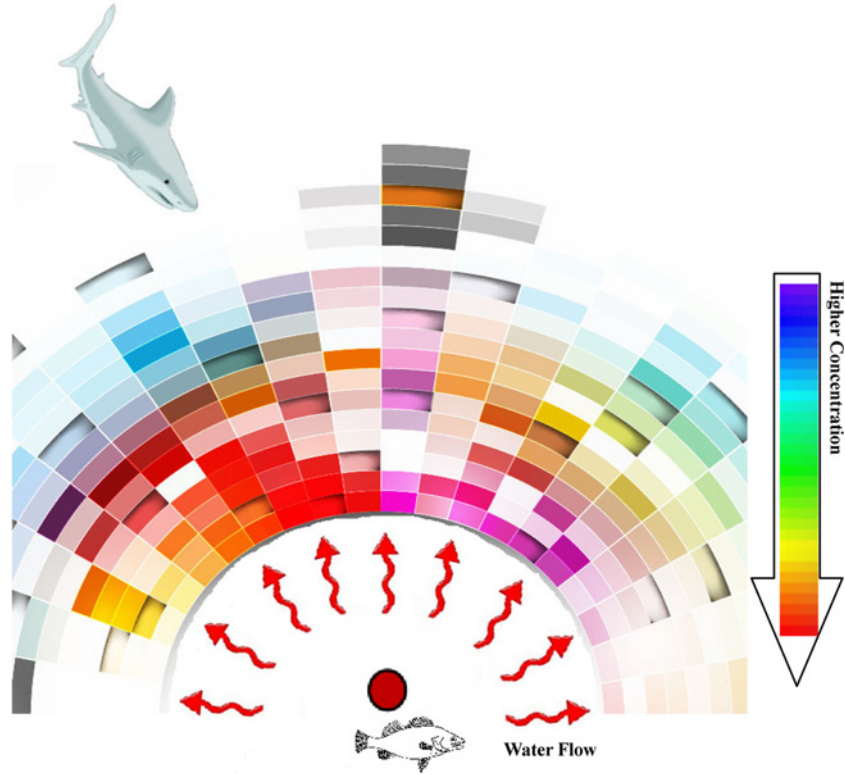
Fig. 3 Schematic illustration of a shark movement to odor source

$$V_{i,j}^{m} = \mu_m . R1 . \frac{\partial(OF)}{\partial x\,j}\bigg|_{x_{i,j}^{m}} \quad (10)$$

$$j = 1, \ldots, ND \,, \; i = 1, \ldots, NP \,, \; m = 1, \ldots, M$$

The current velocity of shark also depends on its previous velocity and so Eq. (10) should be modified as:

$$v_{i,j}^{m} = \mu_m . R1 . \frac{\partial(OF)}{\partial x\,j}\bigg|_{x_{i,j}^{m}} + \alpha_m . R2 . v_{i,j}^{m-1} \quad (11)$$

$$j = 1, \ldots, ND \,, \; i = 1, \ldots, NP \;\; m = 1, \ldots, M$$

where α_m is the inertia constant in the range of $(0,1)$ for stage m. Higher inertia constant leads to more dependency of the current velocity on the previous one. In Eq. (11), $R2$ is a random number with uniform distribution in the range $(0,1)$, similar to $R1$. The shark's velocity in nature has an upper bound. Considering this bound, Eq. (11) becomes as below:

$$\left| v_{i,j}^{m} \right| = M\,in\left[\left| \mu_m . R1 . \frac{\partial(OF)}{\partial x\,j}\bigg|_{x_{i,j}^{m}} + \left| \alpha_m . R2 . v_{i,j}^{m-1} \right| \right|, \left| \gamma_m . v_{i,j}^{m-1} \right| \right] \quad (12)$$

$$j = 1, \ldots, ND \,, \; i = 1, \ldots, NP \;\; m = 1, \ldots, M$$

where γ_m represents the upper bound of current velocity in terms of the previous one. Each element $v_{i,j}^{m}$ of the vector V_{i}^{m} is determined through Eq. (12). The forward movement of shark leads to a new position in the next stage, denoted by Z_{i}^{m+1}, as below:

$$Z_{i}^{m+1} = X_{i}^{m} + V_{i}^{m} . \varDelta t_m \quad i = 1, \ldots, NP \;\; m = 1, \ldots, M \quad (13)$$

where, $\varDelta t_m$ represents the time interval for m^{th} stage. Here, $\varDelta t_m = 1$ is

considered in all stages for the sake of simplicity. Also, the local search of shark can be modeled as:

$$\Gamma_{i}^{m+1,l} = Z_{i}^{m+1} + R3 . Z_{i}^{m+1} \quad (14)$$

$$l = 1, \ldots, L \;\; i = 1, \ldots, NP \;\; m = 1, \ldots, M$$

where, $R3$ is a random number in the interval $(-1,+1)$; L represents number of points in the local search of every stage. Among the points searched in the forward movement and local search, the best one is selected by shark, which is modeled in the SSO algorithm as follows:

$$X_{i}^{k+1} = \operatorname{argmax}\{ OF(Z_{i}^{m+1}), OF(\Gamma_{i}^{m+1,1}), \ldots, OF(\Gamma_{i}^{m+1,L}) \} \quad (15)$$

$$i = 1, \ldots, NP$$

In the selection mechanism of Eq. (18), a maximization problem is assumed. In this paper, the free parameters of the SSO algorithm are set as described in Ref. 19.

Although, the proposed SSO optimization algorithm uses gradient information within the framework of a stochastic search method, gradient operator is combined with random search operators. For instance, in Eq. (12), the gradient operator is multiplied with a random number and then a random part of the previous velocity is added to it through the momentum term. More importantly, a random search operator, i.e. rotational movement, is added to SSO and combined with the forward movement. Thus, if candidate solution traps in a local optimum and cannot proceed through gradient and momentum terms, the random search operator still tries to release it. Even if the random search operator cannot release the candidate solution from local optimum in successive stages, which is a rare case, only one candidate solution is lost. Note that SSO has a population of candidate solutions, which search the solution

space in parallel, another characteristic not seen in a gradient approach. Thus, the other candidate solutions can still search the solution space and find the optimum. This is the general idea of stochastic search techniques that combine analytical and random search operators. For instance, regarding the gradient operator, the stochastic search method of differential evolution uses point-wise gradient (and not analytical gradient used in SSO) and combines it with a random search operator.

3.2 Chaotic SSO

To enhance the search diversity of the SSO, chaotic technology is incorporated into it leading to the proposed CSSO. In the proposed CSSO, the chaotic operator of logistic map is used, which can randomly generate many dissimilar numbers enabling CSSO to cover different areas of the solution space. This operator is as follows:

$$C_{k+1} = 4C_k(1-C_k) \tag{16}$$

where C_k and C_{k+1} are two successively generated numbers by the pendulum chaotic operator in the range of (0,1). To initialize this operator, C_0 is randomly generated in the interval (0,1) such that $C_0 \notin \{0.25, 0.5, 0.75\}$. There are three random number generators in the SSO algorithm including $R1$ and $R2$ related to forward movement and $R3$ pertaining to rotational movement or local search. As $R1$ and $R2$ are in the range of (0,1), the chaotic operator of logistic map is directly used for these random number generators. However, $R3$ is in the range of (-1,+1). Thus, to apply the chaotic operator for $R3$, the following equation is used in addition to Eq. (16):

$$C'_{k+1} = 2C_{k+1} - 1 \tag{17}$$

In other words, $C'_{k+1} \in (-1, +1)$ is used as $R3$. Considering the range of indices in Eq. (12), $ND \times NP$ random numbers $R1$ and $R2$ should be generated. Each of these random number generators includes a sequence of chaotic random numbers by the length of M (i.e. the number of stages), which are generated based on Eq. (16), i.e. $C_k \rightarrow C_{k+1}$. Additionally, considering the range of subscripts in Eq. (14), $NP \times M$ random numbers $R3$ should be produced. Each of these random number generators consists of a sequence of chaotic random numbers by the length of L (i.e. the number of points in local search), which are generated based on Eqs. (16) and (17), i.e. $C_k \rightarrow C'_{k+1}$. Using the chaotic random number generators of $R1$, $R2$ and $R3$, the proposed CSSO can implement a more diverse search compared to SSO, which enhances the chance of finding optimal solution, especially in the complex search spaces, which is the case of training of the NN-based forecasting engine.

4. Numerical Results

The proposed wind power forecast strategy is tested on two real-world case studies. The first one is the Pincher Creek wind farm located in the southern part of Alberta, Canada, and the second one is Sotavento wind farm in Spain. The error criteria used in this paper are Root Mean Square Error (RMSE) and Normalized RMSE (NRMSE):

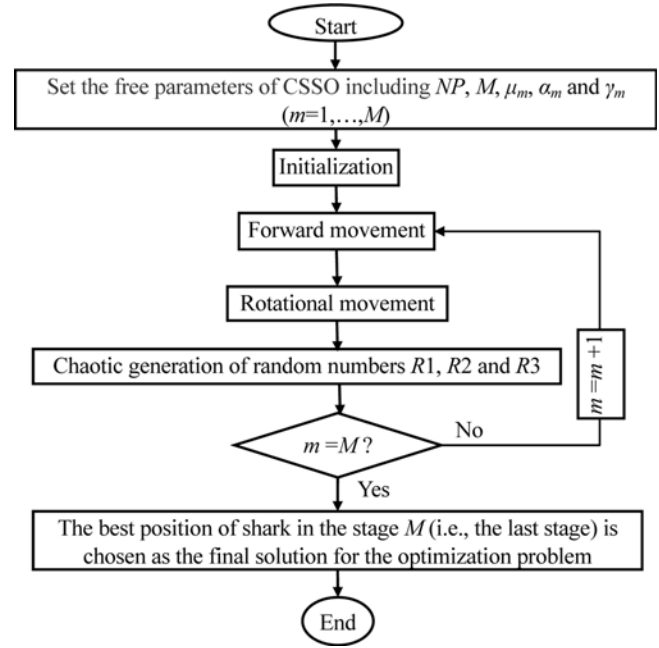$$RMSE = \left[\frac{1}{N}\sum_{t=1}^{N}(S_{ACT(t)}-S_{FOR(t)})^2\right]^{1/2} \tag{18}$$

Fig. 4 Flowchart of chaotic SSO

$$NRMSE = \left[\frac{1}{N}\sum_{t=1}^{N}\left(\frac{S_{ACT(t)}-S_{FOR(t)}}{S_N}\right)^2\right]^{1/2} \tag{19}$$

where $S_{ACT(t)}$ and $S_{FOR(t)}$ indicate the real and predicted values of the signal (here, wind power of the wind farm) for hour t of the forecasting horizon; S_N is the nameplate capacity of the wind farm; N represents the number of hours of the evaluation period. Another error index employed in this paper to measure wind power forecast accuracy is well-known Mean Absolute Percentage Error (MAPE) criterion, which is as follows:

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{|S_{ACT(t)}-S_{FOR(t)}|}{S_{ACT(t)}}\times 100 \tag{20}$$

However, MAPE criterion cannot be directly used to measure wind power forecast error as actual wind power output of a wind farm, i.e. $S_{ACT(t)}$, may become zero in some hours. To remedy this problem a modified version of MAPE criterion, named MMAPE, is used in this paper:

$$MMAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{|S_{ACT(t)}-S_{FOR(t)}|}{S_{AVE-ACT}}\times 100 ,$$
$$S_{AVE-ACT} = \frac{1}{N}\sum_{t=1}^{N}S_{ACR(t)} \tag{21}$$

In other words, the average of wind power values over the evaluation period is used in the denominator of MMAPE to avoid the problem caused by zero values of wind power. The last error criterion considered in this paper is Normalized Mean Absolute Error (NMAE), which is defined as below:

$$NMAE = \frac{1}{N}\sum_{t=1}^{N}\frac{|S_{ACT(t)}-S_{FOR(t)}|}{S_N}\times 100 \tag{22}$$

The five error criteria mentioned above are well-known error measures frequently used in the forecast literature. MAPE is a widely used error criterion as it gives a percentage nature to mean absolute error. However, to evaluate prediction accuracy of a wind power forecast method, it may encounter the problem of infinite error as previously

Table 1 Obtained results for one-hour-ahead wind power prediction in the first test case with the evaluation period of one day

Test day	Error criterion	Persistence[20]	BPNN[20]	RBFNN[20]	ANFIS[20]	NNPSO[20]	WT+ BPNN[20]	WT+ RBFNN[20]	WT+ ANFIS[20]	WT+ NNPSO[20]	Proposed
December 3	MMAPE	10.03	13.62	10.41	14.81	9.54	11.26	8.26	11.08	7.28	7.12
	NMAE	4.18	4.32	4.61	4.75	4.09	4.09	4.19	4.55	3.87	4.02
	NRMSE	5.41	5.73	5.84	6.13	5.38	5.38	5.43	5.91	5.07	4.86
May 4	MMAPE	11.31	12.42	11.07	13.51	11.41	11.41	9.22	11.76	8.73	8.50
	NMAE	4.58	4.88	4.61	4.71	4.51	4.51	4.18	4.39	4.11	4.12
	NRMSE	6.11	6.38	5.89	6.43	6.20	6.20	5.41	6.22	5.83	5.92
July 7	MMAPE	21.58	17.44	16.72	19.30	12.26	12.26	12.39	16.38	11.27	10.12
	NMAE	8.48	7.46	7.32	7.75	5.94	5.94	7.04	7.18	5.29	5.10
	NRMSE	11.25	9.23	8.88	10.16	7.33	7.33	8.36	9.63	7.02	6.76
October 15	MMAPE	14.79	13.93	12.73	12.04	12.82	12.82	14.86	11.08	5.48	5.63
	NMAE	7.48	7.26	7.31	7.76	6.85	6.85	7.08	7.32	6.17	6.08
	NRMSE	9.19	8.79	8.99	9.38	7.43	7.43	8.60	8.93	7.21	6.43
Average	MMAPE	14.43	14.35	12.73	14.91	11.51	11.94	11.18	12.57	8.19	7.84
	NMAE	6.18	5.98	5.96	6.24	5.35	5.35	5.62	5.86	4.86	4.83
	NRMSE	7.99	7.53	7.4	8.02	6.58	6.58	6.95	7.67	6.28	5.99

mentioned. To remedy this problem, the modified version of MAPE or MMAPE is presented, which seems a better choice for assessing wind power forecast error. NMAE gives another measure of mean absolute error, which is normalized in terms of S_N and not $S_{ACT(t)}$ or $S_{AVE-ACT}$ considered in MAPE and MMAPE. On the other hand, RMSE and NMRSE are based on squared error and not absolute error. It is known that squared error gives higher weight to larger errors. NRMSE has a normalized form compared to RMSE. While RMSE measures prediction error in terms of MW and can give an insight to the wind farm operator about the MW power output that is over/under-estimated, NRMSE is a squared error-based criterion, which is dimensionless and can measure the wind power forecast error independent of the wind farm's size. Thus, it cannot be said that which error criterion is the best one as each of them evaluates forecast accuracy from a specific viewpoint. Instead, it is better to give the values of different error criteria for a forecast method to evaluate its prediction accuracy from different aspects and give a better insight to the user about its effectiveness.

4.1 Alberta Test Case

Pincher Creek is a city in the southwest of Alberta, Canada. Its wind farm is considered as the first real-world test case in this paper. For this test case, one-hour-ahead wind power forecasting is considered.[20] A key issue for illustrating the effectiveness of a forecast method is comparing its results with the results obtained from the other methods and published literature figures. However, for the sake of a fair comparison, the same error criteria should be used for all methods; otherwise, the comparison will not be informative. In Table 1, the results obtained from the proposed wind power forecast strategy for this test case are compared with the results of nine other methods in terms of MMAPE, NMA, and NRMSE error criteria. The nine comparative methods of Table 1 include persistence,[20] back propagation NN (BPNN),[20] Radial Basis Function Neural Network (RBFNN),[20] Adaptive Neuro-Fuzzy Inference System (ANFIS)];[20] NN Particle Swarm Optimization (NNPSO),[20] Wavelet Transform (WT) with BPNN,[20] WT+RBFNN,[20] WT+ANFIS,[20] WT+NNPSO.[20] The results reported for these methods in Table 1 are directly quoted from Ref. 20. The same test days of these methods including December 3, July 7, May 4, and October 15 of year 2009 (i.e. one day from each

season) are also considered for the proposed wind power forecast strategy. Thus, the evaluation period in this numerical experiment is one day or 24 hours, i.e. $N=24$ for Eqs. (18)-(22).

Table 1 shows that the proposed wind power forecast strategy provides better average results for the four test days, reported in the last three rows, than all nine other methods in terms of all three error criteria. Even compared to WT+NNPSO, which is the best approach among the nine comparative methods of Table 1, the proposed strategy has 4.3% lower average MMAPE, 0.6% lower average NMAE and 4.6% lower average NRMSE. In addition to obtaining the best average results, the proposed forecast strategy has better results than the nine other methods in most of comparative cases of Table 1.

In the next numerical experiment performed on this test case, the longer evaluation period of three days is considered. The results obtained from this numerical experiment are shown in Table 2. Similar to Table 1, the results reported for the nine comparative methods in Table 2 are directly quoted from Ref. 20. Table 2 shows that the proposed wind power forecast strategy has the best average results among all methods in terms of all three error criteria. Compared to WT+NNPSO, i.e. the best comparative method of Table 2, the proposed strategy leads to 4.8% lower average MMAPE, 0.8% lower average NMAE and 3.6% lower average NRMSE.

To also give a graphical view for the forecast capability of the proposed strategy, its prediction errors in Table 1 are compared with the prediction errors of WT+NNPSO and WT+ANFIS, which are the two best comparative methods of Table 1, in Fig. 5. In each part of this figure, corresponding to one error criterion, the error values of the proposed strategy, WT+NNPSO and WT+ANFIS are connected to each other for creating an error area. The smaller surface of this area means higher prediction accuracy of the associated wind power forecast method. All three parts of Fig. 5 show that the error area of the proposed strategy, indicated by yellow color, has smaller surface than the error areas of WT+NNPSO and WT+ANFIS, indicated by red and grey colors, respectively. Thus, this figure graphically illustrates lower prediction error of the proposed strategy compared to WT+NNPSO and WT+ANFIS. The comparisons of Tables 1 and 2 as well as Fig. 5 clearly represent the wind power prediction capability of the proposed strategy.

Table 2 Obtained results for one-hour-ahead wind power prediction in the first test case with the evaluation period of three days

| Test days	Error criterion	Persistence[20]	BPNN[20]	RBFNN[20]	ANFIS[20]	NNPSO[20]	WT+ BPNN[20]	WT+ RBFNN[20]	WT+ ANFIS[20]	WT+ NNPSO[20]	Proposed
December 3-5	MMAPE	13.27	14.82	11.51	16.29	10.11	13.16	10.09	13.83	8.66	8.12
	NMAE	5.86	4.72	4.76	4.93	4.65	4.54	4.33	4.42	4.11	4.06
	NRMSE	8.11	5.80	5.93	6.24	5.39	5.56	5.62	5.98	5.21	5.08
May 4-6	MMAPE	9.77	12.33	10.89	14.67	11.07	11.34	9.14	12.17	9.32	8.87
	NMAE	5.00	4.74	4.64	4.63	4.28	4.52	4.10	4.42	4.13	4.03
	NRMSE	6.71	6.40	5.82	6.46	6.15	5.94	5.53	6.16	5.64	5.40
July 7-9	MMAPE	17.46	19.87	17.93	20.17	14.79	17.29	15.90	18.63	12.38	11.78
	NMAE	5.36	7.21	7.33	8.16	5.99	7.08	7.10	7.56	5.17	5.24
	NRMSE	8.12	9.82	9.07	9.56	7.83	9.19	8.46	9.12	7.18	7.04
October 15-17	MMAPE	18.29	14.67	12.92	17.94	13.77	13.40	11.85	15.79	12.61	12.12
	NMAE	7.61	7.63	7.83	7.96	7.11	7.38	7.23	7.68	6.55	6.48
	NRMSE	9.90	9.14	9.20	9.42	7.86	8.43	8.74	9.15	7.32	6.92
Average	MMAPE	14.70	15.42	13.31	17.27	12.43	13.80	11.74	15.10	10.74	10.22
	NMAE	5.96	6.075	6.14	6.42	5.51	5.88	5.69	6.02	4.99	4.95
	NRMSE	8.21	7.79	7.50	7.92	6.81	7.28	7.09	7.60	6.34	6.11
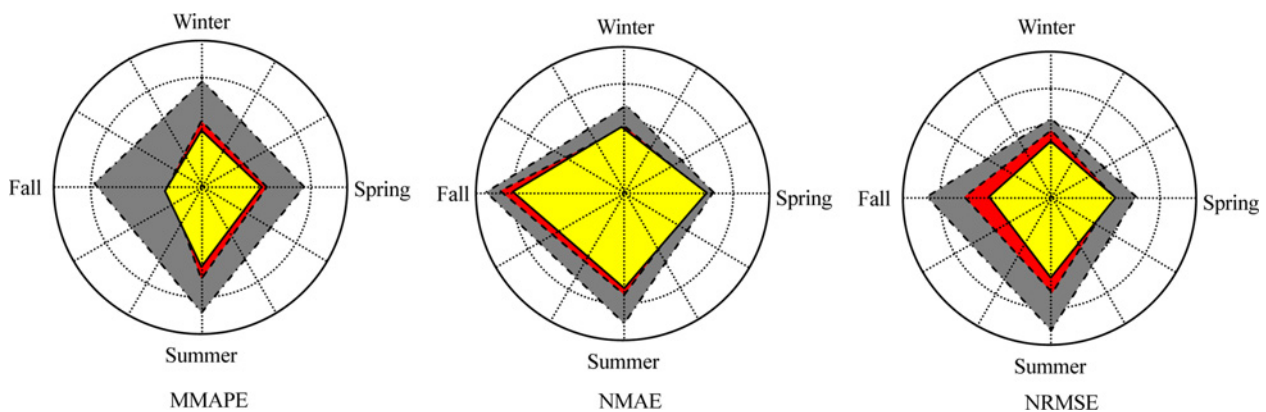
Fig. 5 Comparison of the proposed wind power forecast strategy with WT+NNPSO and WT+ANFIS (Solid: proposed, Dashed: WT+NNPSO, Dashed-Dotted: WT+ANFIS)

Table 3 Obtained numerical results for wind power forecast of Sotavento test case

Test month	Persistence method[22]		Multivariate ARIMA[22]		RBF[22]		MLP[22]		RNN[22]		Proposed	
	RMSE	MMAPE	RMSE	MMAPE	RMSE	MMAPE	RMSE	MMAPE	RMSE	MMAPE	RMSE	MMAPE
April 2010	1.124	35.91	0.843	28.74	0.594	25.08	0.514	22.44	0.463	7.75	0.451	8.12
May 2010	0.848	30.84	0.742	27.21	0.516	18.11	0.618	19.82	0.435	11.43	0.422	10.12
June 2010	0.784	34.33	0.702	28.91	0.593	26.45	0.521	25.24	0.437	16.06	0.434	15.07
July 2010	0.826	36.84	0.691	29.54	0.501	27.75	0.467	26.14	0.376	9.33	0.401	8.59
Average	0.895	34.48	0.744	28.60	0.551	24.35	0.530	23.41	0.428	11.14	0.427	10.47

Table 4 Obtained numerical results for wind speed forecast of Sotavento test case

Test month	Persistence method[22]		Multivariate ARIMA[22]		RBF[22]		MLP[22]		RNN[22]		Proposed	
	RMSE	MMAPE	RMSE	MMAPE	RMSE	MMAPE	RMSE	MMAPE	RMSE	MMAPE	RMSE	MMAPE
April 2010	9.11	99.23	7.92	90.15	6.12	69.34	6.33	70.14	2.13	23.08	2.12	22.21
May 2010	9.21	98.15	7.52	84.34	6.92	76.91	5.38	58.12	1.74	22.23	1.65	21.86
June 2010	9.34	97.69	8.02	85.59	6.48	70.29	6.51	73.39	1.50	21.63	1.55	21.45
July 2010	8.56	98.77	6.34	74.25	6.31	72.34	5.93	66.91	1.51	20.67	1.50	21.01
Average	9.05	98.46	7.45	83.58	6.46	72.22	6.04	67.14	1.72	21.90	1.70	21.63

4.2 Sotavento Test Case

This wind farm is located in Galicia, Spain. It consists of 24 wind turbines of 5 different technologies. The nominal power for this test case is 17.56 MW.[21] In this test case, the proposed wind power forecast strategy is compared with five other methods presented in Ref. 22. The

same test periods of Ref. 22, including four months of April 2010, May 2010, June 2010 and July 2010 are also considered for the proposed strategy. The numerical results obtained for wind power forecast and wind speed forecast are presented in Tables 3 and 4, respectively. In these tables, the proposed strategy is compared with Persistence
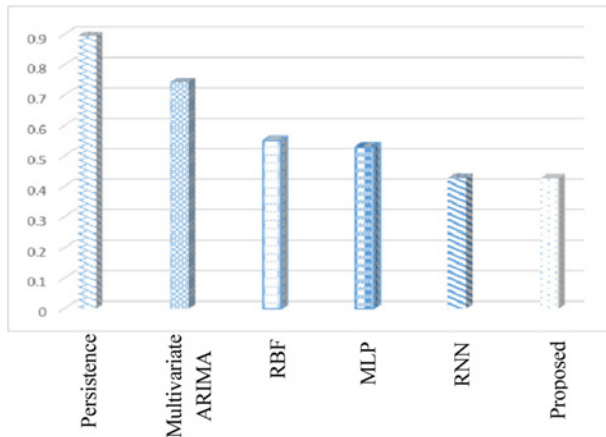
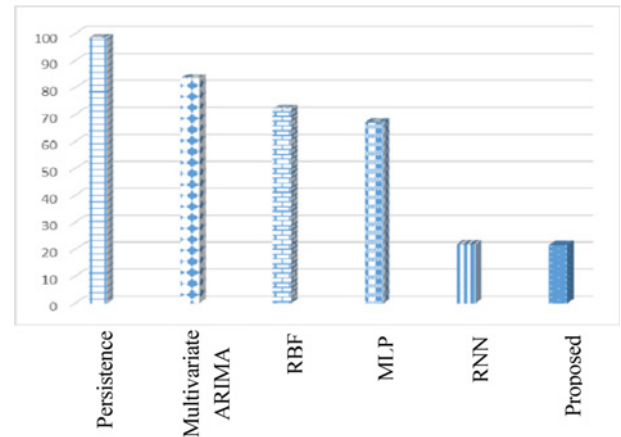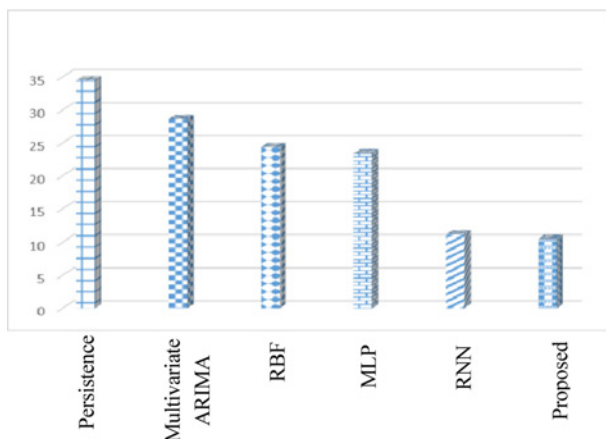Fig. 6 Average RMSE histogram for wind power forecast of the Sotavento test case

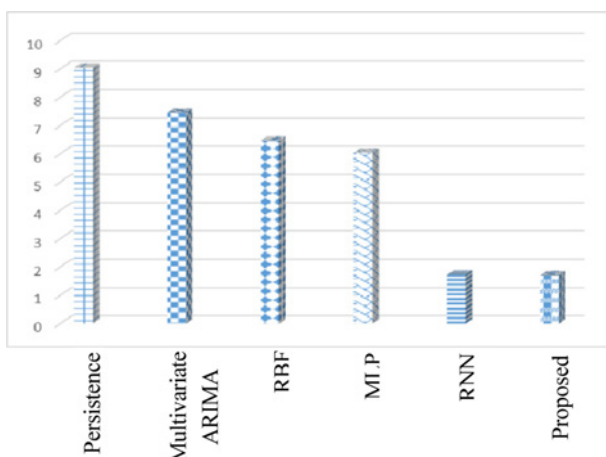Fig. 7 Average MMAPE histogram for wind power forecast of the Sotavento test case

Fig. 8 Average RMSE histogram for wind speed forecast of the Sotavento test case

method, multivariate ARIMA time series, RBF, multi-layer perceptron (MLP) neural network trained by the efficient Levenberg-Marquardt (LM) learning algorithm and Ridgelet Neural Network (RNN). The results of the five comparative methods of Tables 3 and 4 are quoted from Ref. 22. These methods have frequently been used in the other forecast research works Refs. 23-28. Tables 3 and 4 show that the average results of the proposed strategy for both wind power and wind

Fig. 9 Average MMAPE histogram for wind speed forecast of the Sotavento test case
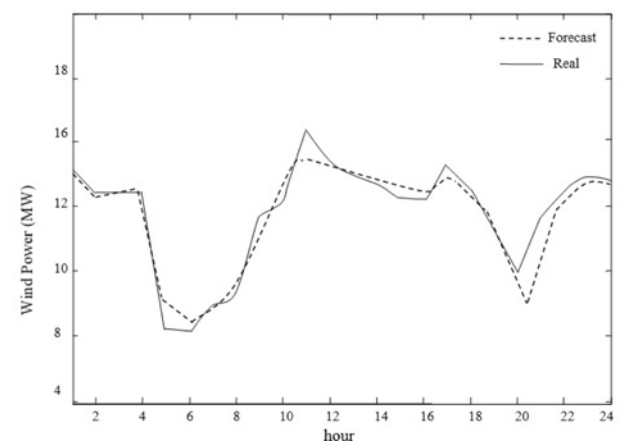

Fig. 10 Curves of real and forecast wind power values for the Sotavento test case in April 18, 2010

Table 5 Obtained results for wind power forecast of the Sotavento test case with different training periods

Test month	30 days		40 days		50 days	
	RMSE	MMAPE	RMSE	MMAPE	RMSE	MMAPE
April 2010	1.006	22.64	0.486	12.98	0.456	8.12
May 2010	1.016	23.14	0.506	12.86	0.406	10.87
June 2010	0.836	20.42	0.569	16.48	0.489	15.47
July 2010	0.859	17.54	0.468	13.68	0.401	9.14
Average	0.929	20.93	0.507	14.00	0.438	10.9

speed prediction are better than the average results of all five other methods in terms of both RMSE and MMAPE error criteria. The average wind power and wind speed prediction results in terms of RMSE and MMAPE are presented in Figs. 6-9, respectively. These figures graphically illustrate lower average errors of the proposed strategy indicating overall better wind power and wind speed forecast performance of the proposed strategy compared to the other methods. In Fig. 10, sample results of the proposed method for wind power prediction of a typical day, i.e. April 18, 2010, including the real curve and forecast curve are shown. This figure shows that the forecast curve reasonably follows the real curve and only small deviations are seen in it.

Selecting appropriate training period is an important factor for the effectiveness of a neural network-based forecast method. A short

training period with small number of training samples can lead to premature convergence of the learning phase and so the neural network cannot correctly learn the input/output mapping function of the forecast process. On the other hand, a long training period may include old historical data, which are irrelevant for training of the neural network. Using such historical data only complicates the learning process for the neural network. To remedy this problem, usually the neural network is run with training periods of different lengths and the training period leading to the best forecast results is selected Ref. 29. In Table 5, three sample results of this study to select an appropriate training period are illustrated. These results are obtained with three training periods including 30 days, 40 days and 50 days prior to the forecast day. It is seen that the training period with the length of 50 days leads to the lowest average errors in terms of both RMSE and MMAPE (reported in the last row of Table 5) for wind power forecast of the Sotavento test case.

The training period of the proposed wind power forecast method in Table 3 is 50 days. Thus, the results of the proposed method in Table 3 are close to the results of Table 5 with 50 days training period. Slight difference between these two sets of results is related to the CSSO part of the proposed wind power forecast method. As any other evolutionary algorithm, CSSO begins from random initial points and thus the results of its different runs are slightly different. This leads to small difference between the results of Table 3 and results of Table 5 with 50 days training period.

5. Conclusions

In this paper, a new wind power forecast strategy, composed of back-propagation NN and evolutionary method, is proposed. Wind power, appeared as a volatile time series, is a complex nonlinear mapping function of many input variables. In the proposed forecast strategy, the prediction performance of back-propagation NN-based forecast engine is improved by means of the evolutionary algorithm of CSSO. CSSO is an enhanced version of SSO, which is a stochastic search method, inspired from shark smell capabilities in nature. CSSO incorporates chaotic technology into SSO to improve its search diversity. CSSO optimizes the initial weight and bias values of the back-propagation NN. Effectiveness of the proposed wind power forecast strategy is extensively compared with 14 other prediction methods on two real-world test cases. These comparisons illustrate wind power forecast capability of the proposed strategy.

REFERENCES

1. Kusiak, A., Zheng, H., and Song, Z., "Short-Term Prediction of Wind Farm Power: A Data Mining Approach," IEEE Transactions on Energy Conversion, Vol. 24, No. 1, pp. 125-136, 2009.

2. Hocaoglu, F. O., Oysal, Y., and Kurban, M., "Missing Wind Data Forecasting with Adaptive Neuro-Fuzzy Inference System," Neural Computing and Applications, Vol. 18, No. 3, pp. 207-212, 2009.

3. Hervás-Martínez, C., Salcedo-Sanz, S., Gutiérrez, P. A., Ortiz-García, E.G., and Prieto, L., "Evolutionary Product Unit Neural Networks for Short-Term Wind Speed Forecasting in Wind Farms," Neural Computing and Applications, Vol. 21, No. 5, pp. 993-1005, 2012.

4. Maqsood, I., Khan, M. R., and Abraham, A., "An Ensemble of Neural Networks for Weather Forecasting," Neural Computing and Applications, Vol. 13, No. 2, pp. 112-122, 2004.

5. Bao, Y., Wang, H., and Wang, B., "Short-Term Wind Power Prediction Using Differential EMD and Relevance Vector Machine," Neural Computing and Applications, Vol. 25, No. 2, pp. 283-289, 2014.

6. Khosravi, A., Nahavandi, S., and Creighton, D., "Prediction Intervals for Short-Term Wind Farm Power Generation Forecasts," IEEE Transactions on Sustainable Energy, Vol. 4, No. 3, pp. 602-610, 2013.

7. Jónsson, T., Pinson, P., Nielsen, H. A., Madsen, H., and Nielsen, T. S., "Forecasting Electricity Spot Prices Accounting for Wind Power Predictions," IEEE Transactions on Sustainable Energy, Vol. 4, No. 1, pp. 210-218, 2013.

8. Amjady, N., Keynia, F., and Zareipour, H., "A New Hybrid Iterative Method for Short-Term Wind Speed Forecasting," European Transactions on Electrical Power, Vol. 21, No. 1, pp. 581-595, 2011.

9. Amjady, N., Keynia, F., and Zareipour, H., "Short-Term Wind Power Forecasting Using Ridgelet Neural Network," Electric Power Systems Research, Vol. 81, No. 12, pp. 2099-2107, 2011.

10. Soman, S. S., Zareipour, H., Malik, O., and Mandal, P., "A Review of Wind Power and Wind Speed Forecasting Methods with Different Time Horizons," in: North American Power Symposium, pp. 1-8, IEEE, 2010.

11. Wang, X., Guo, P., and Huang, X., "A Review of Wind Power Forecasting Models," Energy Procedia, Vol. 12, pp. 770-778, 2011.

12. Bhandari, B., Lee, K.-T., Lee, G.-Y., Cho, Y.-M., and Ahn, S.-H., "Optimization of Hybrid Renewable Energy Power Systems: A Review," Int. J. Precis. Eng. Manuf.-Green Tech., Vol. 2, No. 1, pp. 99-112, 2015.

13. Chang, W.-Y., "A Literature Review of Wind Forecasting Methods," Journal of Power and Energy Engineering, Vol. 2, No. 4, pp. 161-168, 2014.

14. Amjady, N. and Keynia, F., "A New Prediction Strategy for Price Spike Forecasting of Day-Ahead Electricity Markets," Applied Soft Computing, Vol. 11, No. 6, pp. 4246-4256, 2011.

15. Schmitt, M., "On the Complexity of Computing and Learning with Multiplicative Neural Networks," Neural Computation, Vol. 14, No. 2, pp. 241-301, 2002.

16. Wang, W. and Yu, B., "Text Categorization Based on Combination of Modified Back Propagation Neural Network and Latent Semantic Analysis," Neural Computing and Applications, Vol. 18, No. 8, pp. 875-881, 2009.

17. Margaris, A. and Goulianas, K., "Finding All Roots of 2×2 Nonlinear Algebraic Systems Using Back-Propagation Neural Networks," Neural Computing and Applications, Vol. 21, No. 5, pp. 891-904, 2010.

18. Bansal, A., Chen, T., and Zhong, S., "Privacy Preserving Back-Propagation Neural Network Learning over Arbitrarily Partitioned Data," Neural Computing and Applications, Vol. 20, No. 1, pp. 143-150, 2011.

19. Abedinia, O., Amjady, N., and Ghasemi, A., "A New Metaheuristic Algorithm Based on Shark Smell Optimization," Complexity, 2014.

20. Mandal, P., Zareipour, H., and Rosehart, W. D., "Forecasting Aggregated Wind Power Production of Multiple Wind Farms Using Hybrid Wavelet-PSO-NNs," International Journal of Energy Research, Vol. 38, No. 13, pp. 1654-1666, 2014.

21. Sotavento, "Rear Time Data," http://www.sotaventogalicia.com/en/real-time-data/historical (Accessed 26 June 2015)

22. Amjady, N., Keynia, F., and Zareipour, H., "Short-Term Wind Power Forecasting Using Ridgelet Neural Network," Electric Power Systems Research, Vol. 81, No. 12, pp. 2099-2107, 2011.

23. Fan, S., Liao, J. R., Yokoyama, R., Chen, L., and Lee, W.-J., "Forecasting the Wind Generation Using a Two-Stage Network Based on Meteorological Information," IEEE Transactions on Energy Conversion, Vol. 24, No. 2, pp. 474-482, 2009.

24. Gao, X., Sun, Y., and Katayama, S., "Neural Network of Plume and Spatter for Monitoring High-Power Disk Laser Welding," Int. J. Precis. Eng. Manuf.-Green Tech., Vol. 1, No. 4, pp. 293-298, 2014.

25. Costa, A., Crespo, A., Navarro, J., Lizcano, G., Madsen, H., et al., "A Review on the Young History of the Wind Power Short-Term Prediction," Renewable and Sustainable Energy Reviews, Vol. 12, No. 6, pp. 1725-1744, 2008.

26. Lei, M., Shiyan, L., Chuanwen, J., Hongling, L., and Yan, Z., "A Review on the Forecasting of Wind Speed and Generated Power," Renewable and Sustainable Energy Reviews, Vol. 13, No. 4, pp. 915-920, 2009.

27. Damousis, I. G., Alexiadis, M. C., Theocharis, J. B., and Dokopoulos, P. S., "A Fuzzy Model for Wind Speed Prediction and Power Generation in Wind Parks Using Spatial Correlation," IEEE Transactions on Energy Conversion, Vol. 19, No. 2, pp. 352-361, 2004.

28. Kavasseri, R. G. and Seetharaman, K., "Day-Ahead Wind Speed Forecasting Using F-ARIMA Models," Renewable Energy, Vol. 34, No. 5, pp. 1388-1393, 2009.

29. Shahidehpour, M., Yamin, H., and Li, Z., "Market Operations in Electric Power Systems: Forecasting, Scheduling, and Risk Management," New York: IEEE-Wiley, 2002.