Energy Conversion and Management 121 (2016) 232–240

Contents lists available at ScienceDirect

Energy Conversion and Management

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

A novel hybrid approach for predicting wind farm power production based on wavelet transform, hybrid neural networks and imperialist competitive algorithm

Afshin Aghajani, Rasool Kazemzadeh *, Afshin Ebrahimi

Renewable Energy Research Center, Electrical Engineering Faculty, Sahand University of Technology, Tabriz, Iran

A R T I C L E   I N F O

Article history:
Received 24 December 2015
Received in revised form 26 April 2016
Accepted 9 May 2016
Available online 19 May 2016

Keywords:
Neural networks
Imperialist competitive algorithm
Wavelet transform
Wind power forecasting
Forecasting engine

A B S T R A C T

This paper proposes a novel hybrid approach to forecast electric power production in wind farms. Wavelet transform (WT) is employed to filter input data of wind power, while radial basis function (RBF) neural network is utilized for primary prediction. For better predictions the main forecasting engine is comprised of three multilayer perceptron (MLP) neural networks by different learning algorithms of Levenberg–Marquardt (LM), Broyden–Fletcher–Goldfarb–Shanno (BFGS), and Bayesian regularization (BR). Meta-heuristic technique Imperialist Competitive Algorithm (ICA) is used to optimize neural networks' weightings in order to escape from local minima. In the forecast process, the real data of wind farms located in the southern part of Alberta, Canada, are used to train and test the proposed model. The data are a complete set of six meteorological and technical characteristics, including wind speed, wind power, wind direction, temperature, pressure, and air humidity. In order to demonstrate the efficiency of the proposed method, it is compared with several other wind power forecast techniques. Results of optimizations indicate the superiority of the proposed method over the other mentioned techniques; and, forecasting error is remarkably reduced. For instance, the average normalized root mean square error (NRMSE) and average mean absolute percentage error (MAPE) are respectively 11% and 14% lower for the proposed method in 1-h-ahead forecasts over a 24-h period with six types of input than those for the best of the compared models.

1. Introduction

Ever-increasing growth in energy demand throughout the world, fossil fuel resources' reduction [1], and high rate of pollution and the resulting imposed costs have led the electric energy producers to move toward harvesting renewable energy [2]. Among renewable energy types, wind energy is of great importance due to its abundance and relatively low costs [3,4], leading to its enormous utilization in recent years [5]. However, intermittency of wind is the biggest challenge to implementing wind-energy a reliable autonomous source of electric power [2,6–8]. This problem becomes worse when wind power penetration in electric systems increases [9–11]. Undefined supply and its unexpected variations are main obstacles to integrate wind farms into interconnected power systems [12]. In addition, by the introduction of competitive electricity market in various countries, power producers should inform on their produced power several hours ahead to the system operator [1]. Thus, being equipped by accurate predicting tools of power supply [13], this uncertainty will be reduced and wind power's competition place will be elevated in deregulated electricity markets [11,14]. Since the output power from wind plants is highly dependent on environmental condition, the prediction of their output power comes with errors. Obviously, the errors affect system performance [15,16]. More accurate forecasts help system operators plan system operation in a way that absorbing maximum wind energy is possible with minimal complications. This effectively results in lower risks for both producers and the network [17]. Searching through the related literature revealed that huge research, in recent years, has been conducted on predicting power and speed of wind [14,18–21]. The major aim of the reported research is to highly reduce prediction error introduced in significant problems such as wind powers in competitive electricity market [22]. Although the existing methods have made considerable improvements over the years, more accurate and robust wind power forecast methods are still demanded [22]. Predictions are primarily conducted in four time intervals: very short-term,

* Corresponding author.
E-mail addresses: a_aghajani@sut.ac.ir (A. Aghajani), r.kazemzadeh@sut.ac.ir (R. Kazemzadeh), aebrahimi@sut.ac.ir (A. Ebrahimi).

http://dx.doi.org/10.1016/j.enconman.2016.05.024
0196-8904/

short-term, mid-term, and long-term [2,23]. Physical approaches and statistical techniques are the two types of fundamental methods used in literature [18,20,24,25]. The former makes use of physical characteristics of the region where the power plants are installed in order to develop prediction model [26–28]. On the other hand, the latter employs historical data for this purpose. Depending on the model parameters, predictions' accuracy will vary [20,26]. Statistical techniques are divided into two main parts: time-series and artificial neural networks (ANNs) [2]. Time series are effective in short-term predictions [21]. ANNs are among the widely used branches of statistical techniques for predicting wind's power and speed [18,22]. Compared to time-series, ANNs do not need complex mathematical relations for model description, yet they have higher learning capabilities [18]. Due to a nonlinear relation between historical data and wind's speed and power, methods which are able to model nonlinear relations between inputs and outputs should be utilized for precise predictions [29,30]. Neural networks are among the methods which are able to develop nonlinear models through a learning process [29]. In addition to the aforementioned methods, recently, hybrid methods and meta-heuristics as well as novel techniques have been reported in literature [18,29]. The aim of combining forecasting methods is to improve the accuracy of predictions by taking advantage of each method [2]. Since each method itself is sensitive to some conditions, another advantage of combining the methods is to reduce the risk when an unexpected phenomenon occurs [18,31]. For more investigations, one can read wind power forecast in literature [14,18–21]. The main contributions of this research can be summarized as follows:

(1) Proposing a novel hybrid method for short-term prediction of wind farms with high accuracy.
(2) Investigating the effect of six types of different parameters as input data of the algorithm on predictions and comparing it with four types of weather parameters – addition of pressure and air humidity parameters.
(3) Investigating the prediction accuracy of the proposed method in comparison to the other four methods presented in soft computing field.

The rest of the paper is organized as follows. Section 2 describes forecasting methodology, imperialist competitive algorithm, wavelet transform, error measurement criteria, and technical data. In Section 3, forecasting engine is proposed along with its components as well as the proposed algorithm's steps. Section 4 deals with investigation and analysis of results obtained by predictions. Finally, the paper conclusions are given in Section 5.

2. Methodology

This research proposes a novel approach to reduce further predictions error. As statistical techniques utilize historical data, a data mining problem is faced [22]. Data mining is comprised of three parts: data preparation, modeling, and model evaluation. The first step in preparation of data is to extract target data from data sources. In contrast to the other techniques, in this research, a complete set of weather conditions are used as initial data [32]. The next step is data pre-process. Data should be prepared for learning the purpose using various techniques. Therefore, data normalization methods and wavelet transform are used in this research. Modeling refers to a group of processes in which multiple sets of data are combined and analyzed to uncover relationships or patterns. The goal of data modeling is to use past data to inform future efforts. For modeling, a proper technique should be used in a right place. Thus, an appropriate primary predictor should be used to improve the results and help main predictor. Radial networks such as RBF neural networks can be used as the primary predictor. This is because these networks have a special ability to model nonlinear relations and explore local characteristics of the input data. For main predictor, one can use the series combination of several neural networks, including MLP which uses different methods for learning. MLP Neural networks are good at capturing global data trends and modeling nonlinear behaviors. Combination of neural networks of RBF and MLP leads to a consideration of total set of local and global behaviors of the target variables. Meta-heuristic approaches have high exploration capabilities. For further optimization of predictor's engine, the meta-heuristic technique of ICA is used. Considering the above descriptions, the combination of wavelet, RBF, and series neural networks consisting of three types of MLP with different learning strategies along with meta-heuristic algorithm of ICA for prediction are used. In the model evaluation step, using weather parameters of the next day and/or a specific day, the predicted value is evaluated with respect to the real value based on error criteria. These components are described next.

2.1. Wavelet transform

Wavelet transform is a mathematical approach widely used in signal processing applications. This allows to distinguish specific patterns hidden in massive data. Modeling is required when dealing with predictions using time-series and neural networks. Neural networks as general approximators have limited capabilities in approximation of highly nonlinear systems [22]. Wavelet transform has the ability of displaying functions and detecting their local features in time–frequency domain in a simultaneous manner. Low-resolution wavelets and high-resolution wavelets can approximate general behaviors (low frequency) and local behaviors (high frequency) of function, respectively. The use of these features leads to the convenient training along with neural network, precise for modeling highly nonlinear signals [33]. Wavelet transforms are mainly divided into two groups: continuous wavelet transforms (CWT) and discrete wavelet transforms (DWT) [34,35]. The CWT is defined as [36,37]:

$$\mathrm{CWT}_x(a,b) = \frac{1}{\sqrt{|a|}}\int_{-\infty}^{+\infty} \psi^*(t)x(t)dx, \quad a > 0 \tag{1}$$

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), \quad a > 0 \text{ and } -\infty < b < +\infty \tag{2}$$

where $x(t)$ is the signal to be analyzed, $\psi_{a,b}(t)$ is the mother wavelet scaled by a factor a and shifted by a translated parameter b, and $*$ denotes complex conjugate. In a continuous wavelet transform, if scaling and displacement parameters are continuous, CWT will be very slow due to overlapping feature and duplicity of neighbor data. In addition, it will have additional and useless data [12]. Therefore, the mother wavelet can be scaled and translated using certain scales and positions known as DWT. The DWT uses scale and position values based on powers of two, called dyadic dilation and translations, which are obtained by discretized the scaling and translation parameters, denoted as [36,37]

$$\mathrm{DWT}_x(m,n) = 2^{-\left(\frac{m}{2}\right)}\sum_{t=0}^{T-1} x(t)\psi\left(\frac{t-n\cdot 2^m}{2^m}\right) \tag{3}$$

where T is the length of the signal $x(t)$. The scaling and translation parameters are functions of the integer variables m and n, where $a = 2^m$ and $b = n\cdot 2^m$, and t is the discrete time index. Stephane Mallat's multiresolution theory is typically used to employ DWT in related literature [38]. This technique is composed of two major steps: decomposition and reconstruction. Figs. 1 and 2 illustrate related steps in decomposition and reconstruction in this technique

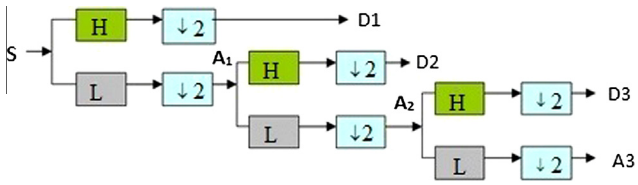

Fig. 1. Wavelet decomposition.

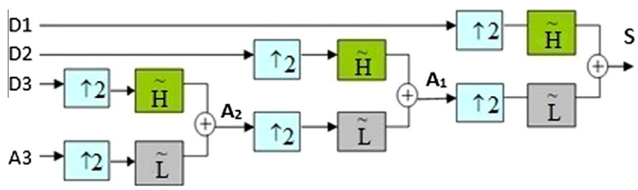

Fig. 2. Wavelet reconstruction.

[12]. Two types of filters are used to decompose the original signal. The high pass filter (H) and low pass filter (L) are used to downsample the original signal as shown in Fig. 1. The decomposed components of the wavelet signal can be assembled back into the original signal without loss of any information. This is called wavelet reconstruction as shown in Fig. 2. Like wavelet decomposition, high pass filter (H) and low pass filter (L) are used, and the decomposed signal is upsampled to obtain the original signal, S [12]. The reconstructed original signal S can be found in the following way:

$$S = A_1 + D_1 = A_2 + D_1 + D_2 = A_3 + D_1 + D_2 + D_3 \tag{4}$$

The approximation (A_1, A_2, A_3) and detail (D_1, D_2, D_3) signals are obtained by downsampling and are only half the length of the original signal. Thus, before reproducing the original signal, it is necessary to reconstruct the approximation and detail coefficients [12].

In decomposition step, the signal is divided into two high and low frequency components. Then, high frequencies are maintained; while, the low frequencies are divided again into two high and low frequency components. High frequencies are called signal details and low frequencies are approximations of the signal. In reconstruction step, actions are reversed in a combination step. There are many types of wavelet basis functions that can be used as a mother wavelet for WTs such as Morlet, Haar, Mexican Hat, and Meyer. Groups of them are called families [39–41]. Among these families, Daubechies have often better results [37,42,43]. In this research, a wavelet function of type Daubechies of order 4 (db4) is used as mother wavelet.

2.2. Imperialist competition algorithm

Imperialist competitive algorithm (ICA) is a new evolutionary algorithm that is inspired by the human's socio-political evolution. This algorithm was first presented by Atashpaz-Gargari and Lucas [44]. Each individual of the population is called a country. The population is divided into two groups, colonies and imperialist states. The competition among imperialists to take possession of the colonies of each other forms the core of this algorithm. In this competition the weak empires collapse gradually and finally there is only one imperialist that all other countries are its colonies. ICA has shown its outstanding ability for the various problems [45–48]. This algorithm is initially started with N colony in which, N_{imp} is the best one (country with the lowest cost) which is selected as imperialisms. In [49,50], ICA pseudocode is described as follows:

1. Selection of the random locations of the function and initialization the empires.
2. Moving the colonies toward their related imperialists (absorption policy or assimilation) according to predetermined assimilation coefficient ($\beta > 1$) and assimilation angle coefficient (γ), which determine the angle and amount of movement.
3. Changing the location of colonies randomly (revolution).
4. It remains in the empire and changes its location relative to imperialist until the cost of colony is less than the imperialist.
5. Uniting the empires with the same conditions.
6. Calculating the total cost of all empires via:

$$\begin{aligned}\text{Total cost of empire} = {} & \text{Cost of imperialist} \\ & + \zeta \times \text{mean(cost of all colonies)}\end{aligned} \tag{5}$$

   where ζ is a constant and mean $(\cdot)$ stands for the average of its arguments.
7. Selecting the weakest colony (colonies) from the weakest empires and put it (them) in one of the empires (colonial competition).
8. Destroying the weak empires.
9. If the preset conditions are satisfied, it will stop, otherwise return to step 2.

2.3. Error measurement criteria

Considering that all predictions naturally have errors, application procedure of error criteria for prediction purpose is highly important [17]. Different criteria have been introduced in literature which can be classified into two general groups [17,51]: first-order, and second-order criteria. One criterion is not sufficient for comparing different methods. Instead, various criteria are required to determine more appropriate technique [17]. First-order criteria have almost similar results. The same thing also occurs for second-order criteria. Second-order criteria have higher sensitivity than first-order ones with respect to the errors. These criteria have no equal amplitudes for comparison purposes. Thus, in order to compare the results, at least one sample from each type should be selected as a representative [17]. These criteria are defined based upon vertical difference of real values. Based on the above definition, two different criteria are used for predictions [12,22]. The mean absolute percentage error (MAPE) criterion is defined as follows [12]:

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\frac{\left|\text{WP}_t^{\text{true}} - \text{WP}_t^{\text{forecast}}\right|}{\overline{\text{WP}_t}^{\text{true},N}} \times 100\% \tag{6}$$

where $N = 24$ for wind power forecasts over a 24-h period, $\text{WP}_t^{\text{true}}$ is the actual wind power in hour t, $\text{WP}_t^{\text{forecast}}$ is the predicted wind power for that hour, and $\overline{\text{WP}_t}^{\text{true},N}$ is the average of actual wind power for all the N test hours, which is given by [12]:

$$\overline{\text{WP}_t}^{\text{true},N} = \frac{1}{N}\sum_{t=1}^{N}\text{WP}_t^{\text{true}} \tag{7}$$

In addition, normalized root mean square error (NRMSE) is also calculated as [22]:

$$\text{NRMSE} = \sqrt{\frac{1}{N}\left(\frac{\text{WP}_t^{\text{true}} - \text{WP}_t^{\text{forecast}}}{\text{WP}_N}\right)^2} \times 100\% \tag{8}$$

where WP_N is the nameplate capacity of wind farm. Note that the total nameplate capacity of Pincher Creek wind farm is 272 MW. MAPE is related to error's first-order and proposes the absolute mean of prediction error. For example, if wind power is predicted, this criterion will be equal to the amount of energy that is deviated from real values in prediction range. In other words, NRMSE criterion is related to second-order error which is changed enormously by high prediction error.

2.4. Technical data

In Table 1 the properties and parameters of the employed optimization algorithms have been brought.

3. The proposed training strategy and forecasting engine

Forecasting engine has two parts [22]: (1) primary predictor and (2) main predictor. Primary predictor is composed of a radial basis function (RBF) neural network which tries to find local characteristics of the input data. Main predictor is constituted of three MLP neural networks using learning algorithms of BR, BFGS, and LM [35] which are in relation with meta-heuristic algorithm, ICA, in order to optimize weightings. Owing to the high exploration capability of the ICA component, an ANN trapped in a local minimum can be released. MLP neural networks have one hidden layer selected from ten neurons by the trial and error method. Fig. 3 depicts main forecasting engine.

Fig. 4 shows schematic diagram of the proposed method for predicting wind power plants. In the following, this method is evaluated in order.

(1) First step is data preparation. As innovative aspect of this research input data are in 6 forms, wind power, wind speed, wind direction, temperature, pressure, and air humidity. Obtaining results from extensive testing of the data types demonstrate that data related to 5 h earlier have the highest impact on wind power data. Therefore, 35 data for 30 days are taken as inputs to train each neural network. In this research, the impact of various combinations of input parameters is evaluated. Used data are taken from wind farms located in southern part of Alberta, Canada.
(2) Second step is primary predictor completed by RBF neural network.
(3) In third step, the output of RBF neural network which is a primary predictor of wind plants' power is added to wind power data. Next, wavelet transform is employed on wind power data and time-series related to wind power are divided into four signals D_1, D_2, D_3, and A_3. For each of the mentioned signals, fourth and fifth steps should be iterated.
(4) In fourth step, with regard to the positive influence of data normalization on neural networks' performance, output data in third step are normalized along with other weather parameters.
(5) Fifth step is related to main predictor. Main predictor is composed of three MLP neural networks with learning algorithms of BR, BFGS, and LM. In each of the networks, once the weightings are found in learning phase, neural network weightings are considered as inputs for meta-heuristic algorithm. When meta-heuristic algorithm gets optimum solutions, its final value is considered as final weightings of neural network. These weightings are considered as the initial weightings for the next neural network. This process is iterated for all three neural networks. Input data in this stage is the outputs of fourth stage. In all these three neural networks, wind power prediction will be desirable. In forecast stage, output of each of these networks is considered along with the others as an input for the next neural network.
(6) After analyzing all of data, output data in fifth step are changed to real value.
(7) In seventh step, wind power's predicted signal is reconstructed after employing reverse wavelet transform in sixth step output.
(8) The above method is run 10 times. And, each time by changing test data, validation and training are performed by the rate (10, 10, 80). Finally, mean predicted values are taken as final results.

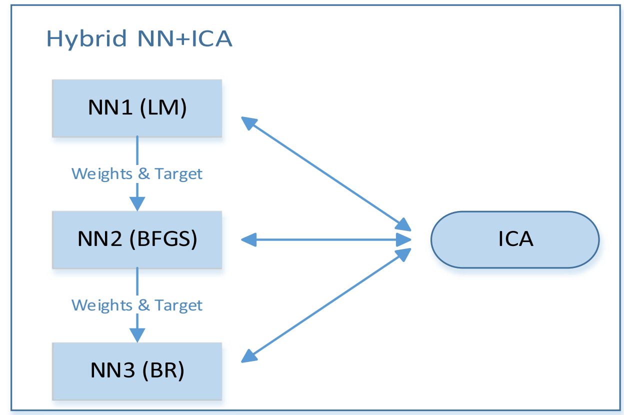

Fig. 3. Main forecasting engine.

In this paper, wind power forecasting has been carried out using two major cases:

- Case I: 1-h-ahead forecasts over 24-h.
- Case II: 1-h-ahead forecasts over 72-h.

The effect of combining various inputs on simulation results will be investigated.

4. Numerical results and discussion

This research uses the data of wind farms located in Pincher Creek, southern part of Alberta, Canada, as the testing system for our forecasting models [52]. There are five wind farms in Pincher

Table 1
The properties and parameters of the employed optimization algorithms.

ANN		ICA		PSO	
Parameters	Value	Parameters	Value	Parameters	Value
Number of neurons in hidden layer	10	Number of initial countries	40	Population size (swarm size)	50
Learning coefficient (η)	0.9	Number of initial imperialists	8	Personal learning coefficient	2
Momentum (α)	0.2	Revolution rate	0.3	Global learning coefficient	2
Activation functions in hidden layer	TANSIG	Assimilation coefficient (β)	2	Inertia weight damping ratio	0.9
Activation functions in output layer	PURELIN	Assimilation angle coefficient (γ)	0.5	Final weight damping ratio	0.4
Number of epochs	1000	ζ	0.02	Max. number of iterations	100

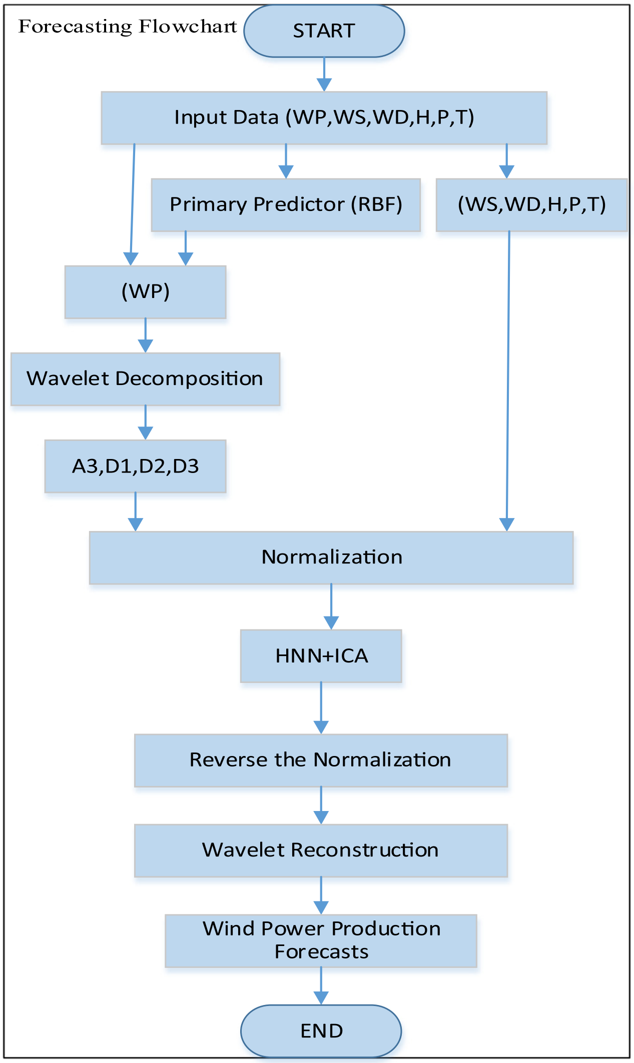

Fig. 4. Schematic diagram of the proposed method.

Creek, namely Castle River (39 MW), Cowley Ridge (38 MW), Kettles Hill (63 MW), Summerview (66 MW), and Summerview-2 (66 MW), making the total capacity of Pincher Creek as 272 MW (p_{max}). This research uses the aggregated wind power of Pincher Creek. In order to examine the impact of pressure and humidity on simulations, simulations are carried out (and compared) in two modes: 4-type conditions and 6-type conditions of input data. Several methods are compared with the proposed technique (RBF + HNN + WT + ICA) including: RBF + HNN + WT + PSO, RBF + HNN + WT, RBF + HNN, HNN + PSO + WT. the simulations were carried out for all these techniques. Table 2 presents simulation results in 1-h-ahead forecasts over a 24-h period with six types of input. As seen, results for four days of different seasons in 2015 are provided. Comparison criteria in this table are MAPE and NRMSE. According to the table, the performance of the proposed method is absolutely better than those of other techniques, leading to a relative reduction of error in all joints. Results obtained in winter and fall are better than other months due to less variations in climatic conditions. Comparing third and fifth columns reveal that meta-heuristic algorithm ICA is better than PSO in finding final weightings of hybrid neural network. In addition, if first and fourth columns are compared, it will quite clear that RBF neural network plays a key role in improving the prediction results. However, if wavelet transform and meta-heuristic algorithm are not utilized, the results given in first column will worsen. Appropriate placement of wavelet transform, inputs, and combination of neural networks along with meta-heuristic algorithm lead to an improvement in the prediction results.

Table 2
Comparison of errors of 1-h-ahead forecasts over a 24-h period with 6 types of input.

Season	Error	Models				
		RBF + HNN	RBF + HNN + WT	RBF + HNN + WT + PSO	HNN + PSO + WT	RBF + HNN + WT + ICA
Winter	MAPE	4.3401	2.4208	2.3908	3.7516	1.8423
	NRMSE	3.0656	1.9176	1.8608	3.7433	1.4041
Spring	MAPE	7.8989	6.6501	3.3021	5.4822	2.6253
	NRMSE	3.824	3.1010	2.4031	3.9214	2.3838
Summer	MAPE	5.6781	3.5198	3.4548	5.1806	3.2563
	NRMSE	3.6397	2.6238	2.4724	4.1519	2.1002
Fall	MAPE	5.2691	4.5889	2.6272	3.7103	2.3314
	NRMSE	4.3413	4.3252	2.2326	3.5461	2.1267

For better comparison, Fig. 5 depicts the contents of the Table 2 in histogram fashion for various techniques.

Obviously, the proposed method outperforms all other techniques. Table 3 provides the simulation results in 1-h-ahead forecasts over a 72-h period with 6 types of input. Considering lengthy time period of predictions, the results bring higher errors than 24-h time period prediction. Thus, it is forecasted to have higher errors in this time-period. Interestingly, the proposed method again gives the best results.

Table 4 illustrates simulation results in 1-h-ahead forecasts over a 24-h period with 4 types of input. Since the effect of air pressure on uncertainties is higher than that of wind speed [18], and considering that it is suggested in literature to take the humidity factor impacts into account on predictions [12], air pressure plus air humidity are added to input data. The impact of various inputs can be evaluated comparing Tables 2 and 4. The effective increase in the number of inputs leads to improvement of results in most cases.

Fig. 6 and 7 illustrate average value of MAPEs and NRMSEs of the four seasons for Tables 2 and 4, respectively. The comparisons reveal the increase in the number of inputs type positively influence the prediction results. For instance, the average NRMSE and average MAPE for the proposed method in 1-h-ahead forecasts over a 24-h period with six types of input are respectively11% and 14% lower than those of the (RBF + HNN + WT + PSO) method.

Accordingly, the proposed method in this case also has the best performance with respect to both error criteria. Figs. 8 and 9 compare the forecasting performance of the proposed model with other hybrid models in terms of MAPE for winter and summer, respectively. From Figs. 8 and 9, it is confirmed that the proposed hybrid intelligent forecasting model performs better even in different cases. For instance, in winter and summer, the average MAPE obtained from the proposed method in 1-h-ahead forecasts over a 24-h period with six types of input are respectively 27% and 8% lower when compared with the proposed method in 1-h-ahead forecasts over a 24-h period with four types of input. Fig. 10 depicts curves of real values, predicted values by the proposed method, and forecasting error of wind power in 1-h-ahead forecasts over a 24-h period with 6 types of input in October 29th of the year 2015.

In one hand, the average computation time required by the hybrid WT + NN + PSO model for short-term (hour-ahead) daily wind power forecasts is around 1–4 min using MATLAB on a PC with 4 GB of RAM and a 2.7-GHz-based processor [12]. On the other hand for the proposed method, the average computation

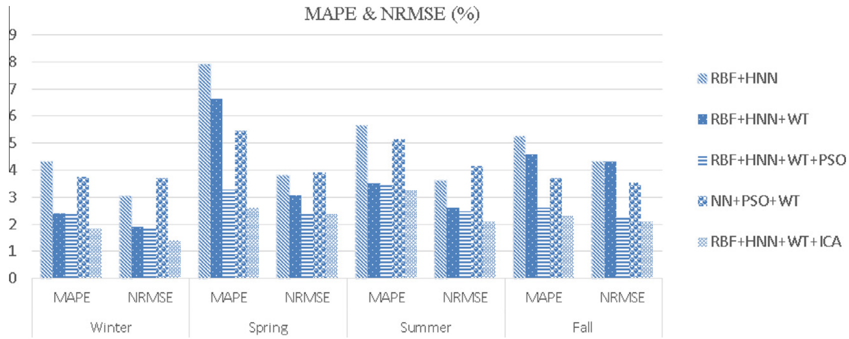

Fig. 5. MAPE & NRMSE for various techniques in 24-h time period with 6 types of input.

Table 3
Comparison of errors of 1-h-ahead forecasts over a 72-h period with 6 types of input.

Season	Error	Models				
		RBF + HNN	RBF + HNN + WT	RBF + HNN + WT + PSO	HNN + PSO + WT	RBF + HNN + WT + ICA
Winter	MAPE	8.1688	3.0217	2.9676	3.9896	2.7557
	NRMSE	5.1754	2.6745	2.3817	3.6743	2.3485
Spring	MAPE	8.5895	7.3339	5.2962	5.9915	4.6249
	NRMSE	5.0630	4.7897	4.3806	4.3123	3.0376
Summer	MAPE	8.9781	4.6985	3.8376	5.4123	3.5740
	NRMSE	6.7993	3.6221	2.8899	4.5637	2.7374
Fall	MAPE	8.7093	5.7106	3.7914	4.0069	3.5107
	NRMSE	5.9347	4.7191	2.6946	3.8201	2.5365

Table 4
Comparison of errors of 1-h-ahead forecasts over a 24-h period with 4 types of input.

Season	Error	Models				
		RBF + HNN	RBF + HNN + WT	RBF + HNN + WT + PSO	HNN + PSO + WT	RBF + HNN + WT + ICA
Winter	MAPE	4.8030	2.9162	2.6510	3.8565	2.3578
	NRMSE	3.9380	2.5674	2.0601	3.5577	1.8463
Spring	MAPE	7.2578	6.9728	3.4360	5.8610	2.7502
	NRMSE	4.7250	4.5722	2.8163	4.2815	2.4330
Summer	MAPE	6.3397	3.8422	3.6035	5.3853	3.5107
	NRMSE	4.0591	2.9380	2.7014	4.2921	2.6948
Fall	MAPE	5.4643	5.2461	2.8571	3.9528	2.7598
	NRMSE	4.7993	4.6895	2.4636	3.7837	2.3552

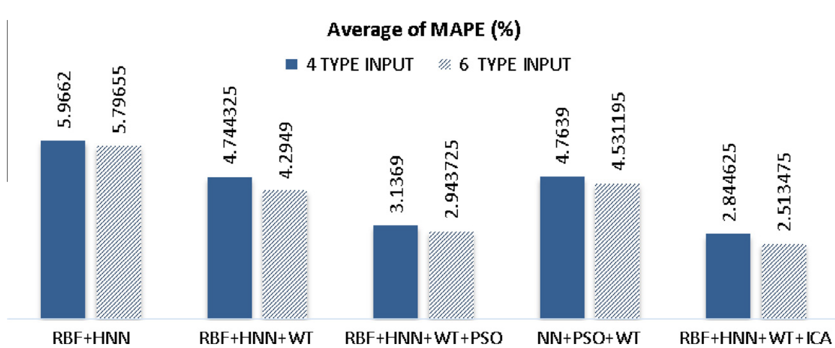

Fig. 6. Average MAPE of the four seasons for Tables 2 and 4.

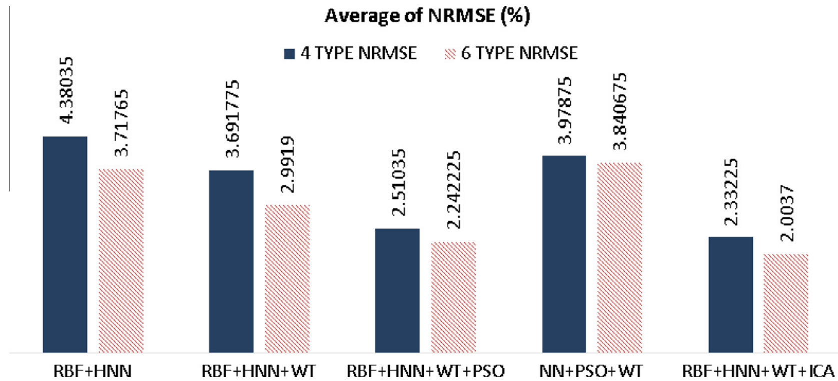

Fig. 7. Average NRMSE of the four seasons for Tables 2 and 4.

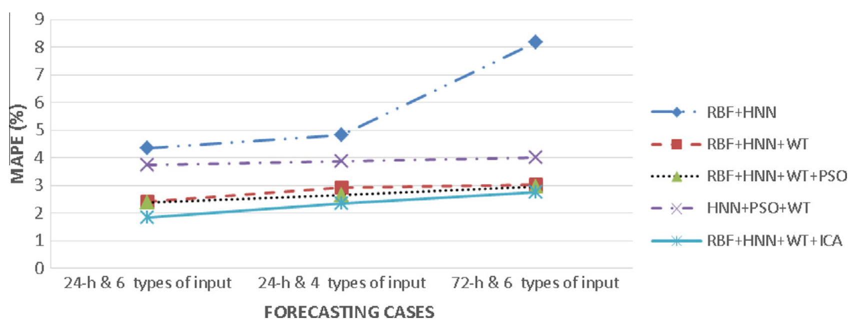

Fig. 8. Comparison of hybrid models for different forecasting cases in winter.

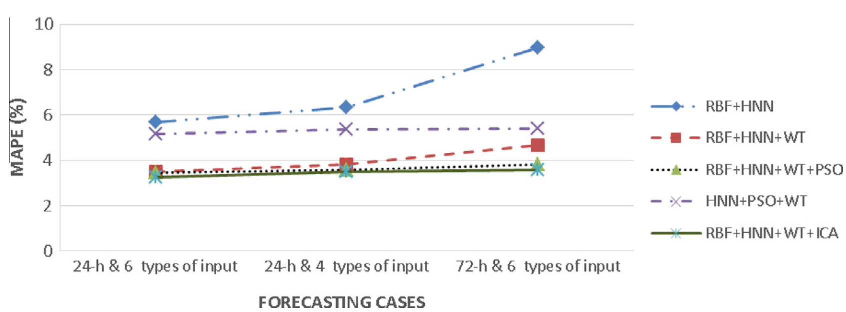

Fig. 9. Comparison of hybrid models for different forecasting cases in summer.

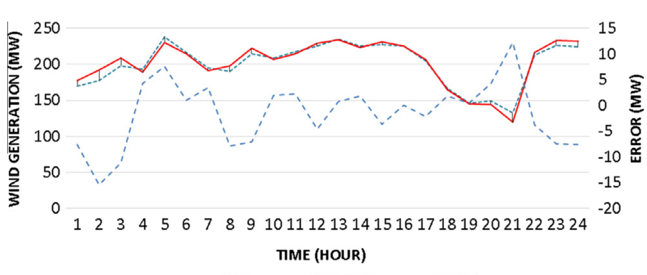

Fig. 10. Curves of real values, forecast values by the proposed method, and forecast errors of wind power in 1-h-ahead forecasts over a 24-h period with 6 types of input.

time is around 1–2 min using MATLAB on a PC with 6 GB of RAM and a 2.1-GHz-based processor. Hence, the proposed approach presents not only better forecasting accuracy, but also an acceptable computation time, which is important for real-life applications [53].

5. Conclusion

The proposed method in this paper is a hybrid technique comprised of wavelet transform, for filtering distortions and noise in wind power signal, and RBF neural network, as primary predictor to find local characteristics of the input data. The main forecasting engine comprised of three MLP neural networks by different learning algorithms along with meta-heuristic algorithm ICA is used for capturing global data trends and appropriate modeling of wind power curve's nonlinear behaviors. In order to demonstrate the efficiency of the proposed method, it is compared with several other wind power forecast techniques. Results of optimizations indicate the *superiority* of the *proposed method* over the other techniques. Results obtained in winter and fall are better than other months due to less variations in climatic conditions. Comparing simulation results reveal that meta-heuristic algorithm ICA is better than PSO in finding final weightings of hybrid neural network. In addition, it is quite clear that RBF neural network plays a key role in improving the prediction results. If wavelet transform and meta-heuristic algorithm are not utilized, the results will worsen. For better comparisons, these simulations were carried out for 1-h-ahead forecasts over a 24-h and 72-h periods. Changing prediction time-period from 24-h to 72-h leads to reduced accuracy of predictions. However, the prediction accuracy of the proposed method was relatively better than that of all other methods. The change of input data from 4 types to 6 types revealed that humidity and pressure are greatly contributed to improvement of simulation results.

Appendix A. Terms and definitions

Here, all terms mentioned in this paper and their definitions are listed in alphabetical order:

A	approximation
BFGS	Broyden–Fletcher–Goldfarb–Shanno
BPNN	backpropagation neural network
BR	Bayesian regularization
CWT	continuous wavelet transform
D	detail
DWT	discrete wavelet transform
H	high pass filter
HNN	hybrid neural network
ICA	imperialist competitive algorithm
L	low pass filter
LM	Levenberg–Marquardt
MAPE	mean absolute percentage error
MLP	multilayer perceptron
NN	neural network
NRMSE	normalized root mean square error
NWP	numerical weather prediction
PSO	particle swarm optimization
RBF	radial basis function
RMSE	root mean square error
SCM	soft computing model
WPF	wind power forecasting
WT	wavelet transform

References

[1] Yuan X, Chen C, Yuan Y, Huang Y, Tan Q. Short-term wind power prediction based on LSSVM–GSA model. Energy Convers Manage 2015;101:393–401.
[2] Soman SS, Zareipour H, Malik O, Mandal P. A review of wind power and wind speed forecasting methods with different time horizons. In: North American Power Symposium (NAPS). IEEE; 2010. p. 1–8.
[3] Montoya FG, Manzano-Agugliaro F, López-Márquez S, Hernández-Escobedo Q, Gil C. Wind turbine selection for wind farm layout using multi-objective evolutionary algorithms. Expert Syst Appl 2014;41:6585–95.
[4] Manzano-Agugliaro F, Alcayde A, Montoya F, Zapata-Sierra A, Gil C. Scientific production of renewable energies worldwide: an overview. Renew Sustain Energy Rev 2013;18:134–43.
[5] Li P, Guan X, Wu J. Aggregated wind power generation probabilistic forecasting based on particle filter. Energy Convers Manage 2015;96:579–87.
[6] Hernandez-Escobedo Q, Manzano-Agugliaro F, Gazquez-Parra JA, Zapata-Sierra A. Is the wind a periodical phenomenon? The case of Mexico. Renew Sustain Energy Rev 2011;15:721–8.
[7] Xu Y, Dong ZY, Xu Z, Meng K, Wong KP. An intelligent dynamic security assessment framework for power systems with wind power. IEEE Trans Ind Inf 2012;8:995–1003.
[8] Hu P, Karki R, Billinton R. Reliability evaluation of generating systems containing wind power and energy storage. IET Gener Transm Distrib 2009;3:783–91.
[9] Kabouris J, Kanellos F. Impacts of large-scale wind penetration on designing and operation of electric power systems. IEEE Trans Sustain Energy 2010;1:107–14.
[10] Pousinho H, Mendes V, Catalão J. A hybrid PSO–ANFIS approach for short-term wind power prediction in Portugal. Energy Convers Manage 2011;52:397–402.
[11] Negnevitsky M, Johnson P, Santoso S. Short term wind power forecasting using hybrid intelligent systems. In: Power Engineering Society General Meeting. IEEE; 2007. p. 1–4.
[12] Mandal P, Zareipour H, Rosehart WD. Forecasting aggregated wind power production of multiple wind farms using hybrid wavelet–PSO–NNs. Int J Energy Res 2014;38:1654–66.
[13] Zhang Y, Wang J, Luo X. Probabilistic wind power forecasting based on logarithmic transformation and boundary kernel. Energy Convers Manage 2015;96:440–51.
[14] Wu Y-K, Hong J-S. A literature review of wind forecasting technology in the world. In: IEEE Lausanne power tech. IEEE; 2007. p. 504–9.
[15] Hernández-Escobedo Q, Saldaña-Flores R, Rodríguez-García E, Manzano-Agugliaro F. Wind energy resource in Northern Mexico. Renew Sustain Energy Rev 2014;32:890–914.
[16] Rohrig K, Lange B. Improvement of the power system reliability by prediction of wind power generation. In: Power engineering society general meeting. IEEE; 2007. p. 1–8.
[17] Shaker H, Zareipour H, Wood D. On error measures in wind forecasting evaluations. In: 26th Annual IEEE Canadian Conference on electrical and computer engineering (CCECE). p. 1–6.
[18] Jung J, Broadwater RP. Current status and future advances for wind speed and power forecasting. Renew Sustain Energy Rev 2014;31:762–77.
[19] Costa A, Crespo A, Navarro J, Lizcano G, Madsen H, Feitosa E. A review on the young history of the wind power short-term prediction. Renew Sustain Energy Rev 2008;12:1725–44.
[20] Giebel G, Brownsword R, Kariniotakis G. The state-of-the-art in short-term prediction of wind power. A literature overview. EU project ANEMOS(ENK5-CT-2002-00665); 2003.
[21] 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;13:915–20.
[22] Amjady N, Keynia F, Zareipour H. Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle swarm optimization. IEEE Trans Sustain Energy 2011;2:265–76.
[23] Zack J. Overview of wind energy generation forecasting Draft report for NY State Energy Research and Development Authority and for NY ISO. Albany (NY, USA): True Wind Solutions LLC; 2003. 17.
[24] Guo Z, Chi D, Wu J, Zhang W. A new wind speed forecasting strategy based on the chaotic time series modelling technique and the Apriori algorithm. Energy Convers Manage 2014;84:140–51.
[25] Sideratos G, Hatziargyriou N. Using radial basis neural networks to estimate wind power production. In: Power engineering society general meeting. IEEE; 2007. p. 1–7.
[26] Ramirez-Rosado IJ, Fernandez-Jimenez LA, Monteiro C, Sousa J, Bessa R. Comparison of two new short-term wind-power forecasting systems. Renew Energy 2009;34:1848–54.
[27] Kariniotakis G, Pinson P, Siebert N, Giebel G, Barthelmie R. The state of the art in short-term prediction of wind power-from an offshore perspective. Proceedings 2004:20–1.
[28] Lange M, Focken U. Physical approach to short-term wind power prediction. Springer; 2006.
[29] Chitsaz H, Amjady N, Zareipour H. Wind power forecast using wavelet neural network trained by improved Clonal selection algorithm. Energy Convers Manage 2015;89:588–98.
[30] Thordarson F, Madsen H, Nielsen HA, Pinson P. Conditional weighted combination of wind power forecasts. Wind Energy 2010;13:751–63.
[31] Hibon M, Evgeniou T. To combine or not to combine: selecting among forecasts and their combinations. Int J Forecast 2005;21:15–24.
[32] Canada's National Climate Archive [Online]. Available: http://climate.weatheroffice.gc.ca/climateData.
[33] Amjady N, Keynia F. Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method. Int J Electr Power Energy Syst 2008;30:533–46.
[34] Conejo AJ, Plazas M, Espinola R, Molina AB. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Trans Power Syst 2005;20:1035–42.
[35] Amjady N, Daraeepour A, Keynia F. Day-ahead electricity price forecasting by modified relief algorithm and hybrid neural network. IET Gener Transm Distrib 2010;4:432–44.
[36] Amjady N, Keynia F. Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm. Energy 2009;34:46–57.
[37] Catalao J, Pousinho H, Mendes V. Hybrid wavelet-PSO-ANFIS approach for short-term wind power forecasting in Portugal. IEEE Trans Sustain Energy 2011;2:50–9.
[38] Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 1989;11:674–93.
[39] Lei C, Ran L. Short-term wind speed forecasting model for wind farm based on wavelet decomposition. In: DRPT 2008 third international conference on electric utility deregulation and restructuring and power technologies. IEEE; 2008. p. 2525–9.
[40] Pandey AS, Singh D, Sinha SK. Intelligent hybrid wavelet models for short-term load forecasting. IEEE Trans Power Syst 2010;25:1266–73.
[41] Chuanan Y, Yongchang Y. A hybrid model to forecast wind speed based on wavelet and neural network. In: International conference on control, automation and systems engineering (CASE). IEEE; 2011. p. 1–4.
[42] Damousis IG, Alexiadis MC, Theocharis JB, Dokopoulos PS. A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. IEEE Trans Energy Convers 2004;19:352–61.
[43] Haque AU, Mandal P, Meng J, Srivastava AK, Tseng T-L, Senjyu T. A novel hybrid approach based on wavelet transform and fuzzy ARTMAP networks for predicting wind farm power production. IEEE Trans Ind Appl 2013;49:2253–61.

[44] Atashpaz-Gargari E, Lucas C. Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition. In: CEC 2007 IEEE congress on evolutionary computation. IEEE; 2007. p. 4661–7.
[45] Abdechiri M, Faez K, Bahrami H. Neural network learning based on chaotic imperialist competitive algorithm. In: 2nd International workshop on intelligent systems and applications (ISA). IEEE; 2007. p. 4661–7.
[46] Niknam T, Fard ET, Pourjafarian N, Rousta A. An efficient hybrid algorithm based on modified imperialist competitive algorithm and K-means for data clustering. Eng Appl Artif Intell 2011;24:306–17.
[47] Jolai F, Sangari MS, Babaie M. Pareto simulated annealing and colonial competitive algorithm to solve an offline scheduling problem with rejection. Proc Inst Mech Eng, Part B: J Eng Manuf 2010;224:1119–31.
[48] Lucas C, Nasiri-Gheidari Z, Tootoonchian F. Application of an imperialist competitive algorithm to the design of a linear induction motor. Energy Convers Manage 2010;51:1407–11.
[49] Mahmoudi MT, Forouzideh N, Lucas C, Taghiyareh F. Artificial neural network weights optimization based on imperialist competitive algorithm. In: 7th International conference on computer science and information technologies (CSIT'09), Yerevan. p. 244–7.
[50] Bigdeli N, Afshar K, Gazafroudi AS, Ramandi MY. A comparative study of optimal hybrid methods for wind power prediction in wind farm of Alberta, Canada. Renew Sustain Energy Rev 2013;27:20–9.
[51] Madsen H, Pinson P, Kariniotakis G, Nielsen HA, Nielsen T. Standardizing the performance evaluation of shortterm wind power prediction models. Wind Eng 2005;29:475–89.
[52] AESO. Wind power integration. [Online]. Available: http://www.aeso.ca/gridoperations/20544.html.
[53] Mathworks. URL:http://www.mathworks.com.