Applied Soft Computing 13 (2013) 3225–3233

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

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

Short-term wind speed forecasting based on a hybrid model

Wenyu Zhang[a,b,*], Jujie Wang[a,b], Jianzhou Wang[c], Zengbao Zhao[a,b], Meng Tian[a,b]

[a] Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China
[b] Key Laboratory of Arid Climatic Change and Reducing Disaster of Gansu Province, Lanzhou 730000, China
[c] School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, China

ARTICLE INFO

Article history:
Received 29 March 2012
Received in revised form 16 January 2013
Accepted 14 February 2013
Available online 4 March 2013

Keywords:
Forecasting
Wind speed
Wavelet transform
Seasonal adjustment
RBF neural networks

ABSTRACT

Wind power is currently one of the types of renewable energy with a large generation capacity. However, operation of wind power generation is very challenging because of the intermittent and stochastic nature of the wind speed. Wind speed forecasting is a very important part of wind parks management and the integration of wind power into electricity grids. As an artificial intelligence algorithm, radial basis function neural network (RBFNN) has been successfully applied into solving forecasting problems. In this paper, a novel approach named WTT–SAM–RBFNN for short-term wind speed forecasting is proposed by applying wavelet transform technique (WTT) into hybrid model which hybrids the seasonal adjustment method (SAM) and the RBFNN. Real data sets of wind speed in Northwest China are used to evaluate the forecasting accuracy of the proposed approach. To avoid the randomness caused by the RBFNN model or the RBFNN part of the hybrid model, all simulations in this study are repeated 30 times to get the average. Numerical results show that the WTT–SAM–RBFNN outperforms the persistence method (PM), multilayer perceptron neural network (MLP), RBFNN, hybrid SAM and RBFNN (SAM–RBFNN), and hybrid WTT and RBFNN (WTT–RBFNN). It is concluded that the proposed approach is an effective way to improve the prediction accuracy.

1. Introduction

In recent years, due to global environmental pollution issues, renewable energy (such as wind, solar, geothermal, biomass, tidal, and hydropower) has received increasing attention. Wind power is one of the cleanest renewable energy sources that produce no greenhouse gases, has no effect on climate change, and produces little environmental impacts, and the energy generated from the wind has been well recognized as environmentally friendly, socially beneficial, and economically competitive for many applications [1]. However, the power generation from the wind has been plagued by the intermittent and stochastic nature of wind source, and thus it is still a less reliable source and difficult to be integrated into power grid systems [2]. This problem can be significantly mitigated if the operation of wind farm can be controlled based on the accurate information of dynamic wind speed forecasting. Accurate short-term wind speed forecasting can reduce voltage and frequency fluctuations due to variation in wind power and unacceptable shocks in the conventional power units caused by a sudden cut-off of wind power resulting from excessive wind speeds [3,4]. As a result, accurate forecasting of the short-term wind speed is a critical issue for improving both the reliability of a wind power generation system and the integration of wind energy into the power system [5].

Short-term wind speed forecasting can be made in the order of several days and also from minutes to hours [6]. Usually, hourly forecasts of expected winds are helpful in dispatching decision-making, daily forecasts of hourly winds are useful for the load scheduling strategy, and weekly forecasts of day-to-day winds greatly facilitate maintenance scheduling [2,7]. Many methods on the short-term wind speed forecasting have been developed over the past two decades. Generally, these methods used in the literature can be divided into two categories: statistical models and artificial intelligence models (AI). Statistical models are identical to the direct random time-series model, including auto regressive (AR), and auto regressive integrated moving average (ARIMA) models [9]. They do well in short-term forecasts, and have been tested in short-term wind speed forecasts [3,8,10–14]. However, they are not perfect in forecasting. First, most of statistical models assume that the wind speed data is normally distributed, but it is well known that wind speed series is not a normally distributed [3]. Second, the intermittent and stochastic characteristic of wind speed series need more complex functions for capturing the nonlinear relations, but these models are based on the assumption that a linear correlation structure exists among time series values [15]. As a result,

* Corresponding author at: College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China. Tel.: +86 931 8911031; fax: +86 931 8914278.
E-mail address: zhangwy@lzu.edu.cn (W. Zhang).

1568-4946/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2013.02.016

the wind speed series is difficult to be predicted accurately. To overcome this limitation of statistical approaches, the AI models, mainly including artificial neural networks (ANNS), have attracted more attention for accurate short-term wind speed forecasting and have also been determined to be more accurate as compared to traditional statistical models [4,8,16–19,11,20–25].

ANNs are data-driven and non-parametric models, and can be useful techniques for wind speed forecasting due to their ability to capture subtle functional relationships among the empirical data even though the underlying relationships are unknown or hard to describe. As a special neural network, radial basis function neural networks (RBFNNs) have been extensively studied by researchers in nonlinear identification and time series forecasting areas [26–29]. They have the ability to rapidly learn complex patterns and tendencies presented in data, with fast adaptation to changes. It has been proved that a RBFNN can approximate arbitrarily well any multivariate continuous function on a compact domain if a sufficient number of radial basis function units are given [30,31]. It motivates this study of using RBFNN for the short-term wind speed forecasting in this paper.

Wind speed series are non-stationary and highly-noisy due to wind being a weather driven renewable resource which mainly depends on the climate system. Forecasting the wind speed data with the noisy directly is usually subject to large errors. The wavelet transform technique (WTT) is an essential tool for data preprocessing and has been widely used in de-noising and extracting the basic characteristics from the non-stationary time series [32]. For this reason, this WTT is applied to the wind speed series for de-noising in the paper. On the other hand, the seasonal and trend variations are the two most commonly encountered phenomena in the wind speed series. The seasonal component information is often neglected in most wind speed series forecasting, whose fluctuation causes large deviation to the forecasting. According to Zhang and Qi [33], de-seasonalization can dramatically reduce forecasting errors in many seasonal time series. The seasonal adjustment method (SAM) which can separate the seasonal component and trend component from the time series can help estimate the trend and make short-term forecasting more efficient, so the SAM is applied for the preprocessing of the wind speed data in this paper.

The WTT can effectively remove the useless information in wind speed series. The SAM is mainly applied to forecast the seasonal component and the RBFNN is adopted to forecast the trend component in the wind speed datasets. Considering the actual features of wind speed series, in this paper, a novel approach named WTT–SAM–RBFNN for short-term wind speed forecasting is proposed by applying WTT into hybrid model which hybrids the SAM and RBFNN. In such a model, the raw wind speed data are first decomposed into an approximate part associated with low frequencies and a detail part associated with high frequencies by the WTT. Second, the SAM separates the seasonal component and trend component from the low-frequency signal series, and simulates and predicts the seasonal component. Third, the trend component is predicted by the RBFNN, and the forecasting value of the raw wind speed data is got by multiplying the seasonal index to the forecasting value of the trend component. Finally, the mean hourly wind speed data about one month in Wuwei and Minqin in China are used as illustrative examples to evaluate the forecasting accuracy of the proposed approach. The empirical results also clearly show that the WTT–SAM–RBFNN outperforms the persistence method (PM), multilayer perceptron neural network (MLP), RBFNN, hybrid SAM and RBFNN (SAM–RBFNN), and hybrid WTT and RBFNN (WTT–RBFNN). It is concluded that the proposed approach is an effective way to improve the forecasting accuracy.

The rest of this paper is organized as follows. Section 2 presents the WTT–SAM–RBFNN approach for forecasting the wind speed. Section 3 provides the evaluation criteria and presents the numerical results from two real world cases study. Finally, Section 4 outlines the conclusions.

2. Proposed approach

The proposed WTT–SAM–RBFNN approach, which applies the WTT into a hybrid model which hybrids SAM and RBFNN, is proposed for forecasting short-term wind speed in this paper. The algorithm is described as follows and the flowchart is shown in Fig. 1. The methods used in the WTT–SAM–RBFNN approach are briefly introduced in the following sections.

Step 1: Wavelet de-noising. The raw wind speed data series are decomposed into a low-frequency component and a high-frequency component by the WTT. The low-frequency component represents the main features of the raw data series, and the high-frequency component is often called the noise signal. The idea of this step is to extract the main characteristics and remove the random disturbance from the raw data series.
Step 2: Removing the seasonal component. This step decomposes low-frequency component into seasonal component and trend component by the SAM, and the seasonal indices $I(s)\,(s=1, 2, \ldots, m)$ are calculated. In this paper, $m=24$ is the sample number in every cycle.
Step 3: Capturing the pattern of the trend component. The trend component is divided into two subsets: training set $T_1(i)\,(i=1, 2, \ldots, n_1)$ and testing set $T_2(i)\,(i=1, 2, \ldots, n_2)$. The RBFNN is mainly applied to capture the complex pattern from the trend component. The RBFNN model is constructed by training set $T_1(i)$, and the fitting value of the model is denoted as $R(i)\,(i=1, 2, \ldots, n_1)$. Here, n_1 and n_2 are the sample numbers in the training set and testing set respectively.
Step 4: Fitting the raw wind speed series. The fitting value of the raw wind speed data series is got by multiplying the seasonal index $I(s)\,(s=1, 2, \ldots, m)$ to the fitting value of the trend component $R(i)$. That is to say $F(i)=R(i)\times I(s)$ $(i=1, 2, \ldots, n_1\,; s=1, 2, \ldots, m)$. Here, $F(i)$ is the fitting value of the raw wind speed series.
Step 5: Validation. The proposed approach is validated by the testing datasets $T_2(i)$, and make forecasts.

2.1. Wavelet transform and de-noising

The wavelet transform technique (WTT) is an essential tool for data pre-processing and has been widely used in de-noising and extracting the basic characteristics from the non-stationary time series [32]. In this section, a brief summary of WWT is presented.

WTT may be regarded as a special type of Fourier transform at multiple scales, and can be divided in two categories: continuous wavelet transforms (CWT) and discrete wavelet transforms (DWT). The CWT is defined as the convolution of a time series $x(t)$ with a wavelet function $w(t)$ [34]:

$$CWT_x^{\psi}(b,a)=\phi_x^{\psi}(b,a)=\frac{1}{\sqrt{|a|}}\int x(t)\cdot\psi^{\bullet}\left(\frac{t-b}{a}\right)dt \tag{1}$$

where a is a scale parameter, b is the translational parameter, and $\bullet$ is the complex conjugate of $\psi(t)$. Let $a=1/2^s$ and $b=k/2^s$, where s and k belong to the integer set Z. The CWT of $x(t)$ is a number at $(k/2^s, 1/2^s)$ on the time-scale plane. It represents the correlation between $x(t)$ and $\psi^{\bullet}(t)$ at that time-scale point.

A discrete version of Eq. (1) is thus obtained as

$$DWT_x^{\psi}(k,s)=\phi_x^{\psi}\left(\frac{k}{2^s},\frac{1}{2^s}\right)=\int_{\infty}^{\infty}x(t)\cdot\psi^{\bullet}\left(\frac{t-k/2^s}{1/2^s}\right)dt \tag{2}$$

which separates the signal into components at various scales corresponding to successive frequencies. Note that DWT corresponds

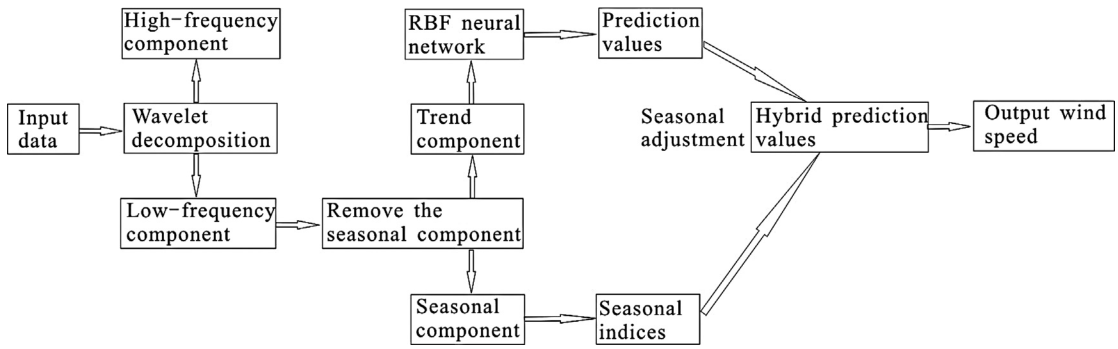

Fig. 1. The flowchart of the proposed WTT–SAM–RBFNN approach.

to the multi-resolution approximation expressions for analysis of a signal in many frequency bands (or at many scales). In practice, multi-resolution analysis is carried out by starting with two channel filter banks composed of a low-pass and a high-pass filter, and then each filter bank is sampled at a half rate of the previous frequency. The number of steps of this decomposition procedure will depend on the length of data. The down sampling procedure keeps the scaling parameter constant (1/2) throughout successive wavelet transforms [35,36].

Wavelet de-noising technique based on DWT is a noise removal technique and has been applied in the fields of image processing, signal processing and time series analysis [37–41,46]. In the proposed model, the wavelet de-noising based on hard thresholding method is first used to filter out the noise from the wind speed. Then the SAM and RBFNN utilize the filtered wind speed time series to build the forecasting model. When using wavelet de-noising, two parameters must be settled. They are wavelet basis function and levels of decomposition. In the wavelet de-noising literature, a lot of wavelet functions are used. According to the difference of resolution capability and efficiency, every wavelet function has the corresponding application field. The daubechies wavelet filters of order 3 (db3) has the high computing efficiency and is widely adopted in time series analysis. In this study, the db3 is adopted in the wavelet de-noising process. In addition, considering the characteristics of the experimental data, after pre-testing different levels with db3, level 1 seems to work best with most sets of the data of this study. For more detail information about wavelet transform and wavelet de-noising technique, please refer to [40,42–46].

2.2. Seasonal adjustment method (SAM)

The seasonal component and the trend component actually coexist in the wind speed series. The SAM can be applied to separate the seasonal item from the time series, and it can help estimate the trend and make short-term forecasting more efficient. In general, seasonal adjustment can be divided into the addition model and multiplication model. According to Zhang and Qi [33], the classical multiplicative decomposition model is appropriate for many time series with seasonal variations. As a result, we adopted the multiplicative decomposition models in our case study to forecast wind speed. The basic idea of the multiplicative decomposition models can be described as follows [47]:

Assuming that the wind speed at time t can be expressed as

$$x_t = f(t) I_j \tag{3}$$

where $f(t)$ and I_j represent the trend component and seasonal component respectively, then, the seasonal index I_j can be obtained by

$$I_j = \frac{x_t}{f(t)} \tag{4}$$

Because the trend component was unknown, we used the average of x_i in each cycle as its approximation.

If we rearrange the dataset $x_1, x_2, \ldots, x_T$ to be $x_{11}, x_{12}, \ldots, x_{1l}, x_{21}, x_{22}, \ldots, x_{2l}, \ldots, x_{k1}, x_{k2}, \ldots, x_{kl}, \ldots, x_{m1}, x_{m2}, \ldots, x_{ml} (k = 1, 2, \ldots, m; s = 1, 2, \ldots, l)$, where m and l represent the number of cycles and the number of data items in each cycle, respectively, and $T = m \times l$, then the average can be derived by

$$\bar{x}_k = \frac{x_{k1} + x_{k2} + \cdots x_{kl}}{l} \quad (k = 1, 2, \ldots, m) \tag{5}$$

Normalizing the data items x_{ks}, we have

$$I_{ks} = \frac{x_{ks}}{\bar{x}_k} \quad (k = 1, 2, \ldots, m;\ s = 1, 2, \ldots, l) \tag{6}$$

I_j is defined as

$$I_j = \frac{I_{1j} + I_{2j} + \cdots I_{mj}}{m} \quad (j = 1, 2, \ldots, l) \tag{7}$$

This definition of I_j conforms to the normalization, as demonstrated by

$$\sum_{j=1}^{l} I_j = \frac{1}{m} \sum_{k=1}^{m} \sum_{s=1}^{l} I_{ks} = \frac{1}{m} \sum_{k=1}^{m} \left(\sum_{s=1}^{l} \frac{x_{ks}}{\bar{x}_k} \right) = \frac{1}{m} \sum_{k=1}^{m} l = l$$

Using the value of I_s, the series without the impact of the seasonal component can be obtained by

$$x'_{ks} = \frac{x_{ks}}{I_s} \quad (k = 1, 2, \ldots, m;\ s = 1, 2, \ldots, l) \tag{8}$$

If we rearrange the data items $x'_{11}, x'_{21}, \ldots, x'_{1l}; \ldots, x'_{m1}, x'_{m1}, \ldots, x'_{ml}$ back to $x'_1, x'_2, \ldots, x'_T$, we then obtain the new data series without the seasonal component.

2.3. Radial basis function neural networks (RBFNNs)

Radial basis function is a special class of functions, their characteristic feature is that response decreases, or increases, monotonically with distance from a center point. It has been shown that the RBFNN are able to approximate any reasonable continuous function mapping with a satisfactory level of accuracy and are more useful than the back propagation neural networks by some investigators [48,49]. It was decided to see if the RBFNN work well for this application as well.

Normally, a RBFNN is generally consists of three layers, input layer, hidden layer, and output layer. The basic architecture of an RBFNN is shown in Fig. 2. The input layer feeds the input data to each of the nodes of the hidden layer. Each node of the hidden layer represents a data cluster and differs greatly from other neural networks. Each node of the output layer sums the outputs of the previous layer and is to yield a final output value [50,51]. The calculated process can be described as follows.

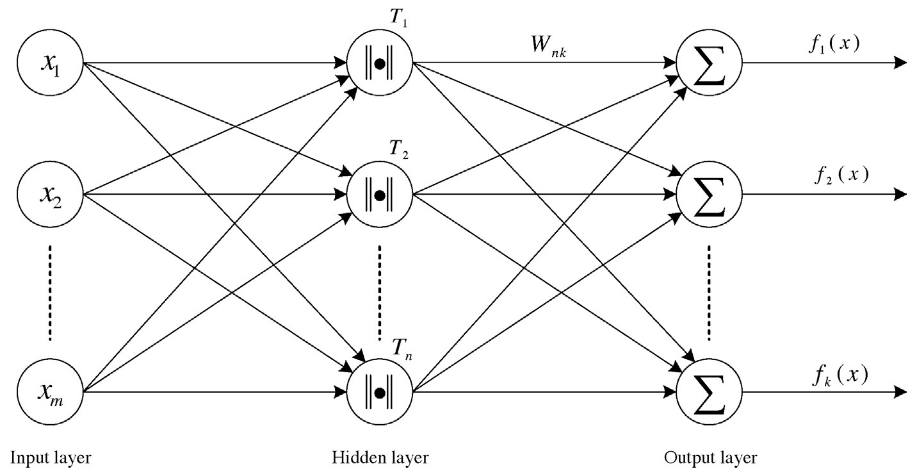

Fig. 2. The structure of RBF neural network.

The output of the RBFNN is deduced by

$$y = f(x) = \sum_{k=1}^{N} w_k \phi_k(\left\| x - c_k \right\|, \sigma_k) \tag{9}$$

where y is the actual network output, $x \in R^{m\times1}$ is an input vector signal, with individual vector components given by x_j, for $j = 1, 2, \ldots, m$, that is, $x = [x_1, x_2, \ldots, x_m,]^T \in R^{m\times1}$. $w = [w_1, w_2, \ldots, w_N,]^T \in R^{N\times1}$ is the vector of the weights in the output layer, N is the number of neurons in the hidden layer, and $\phi_k(\bullet) : R^{m\times1} \to R$ is the basis function of the network. $c_k = [c_{k1}, c_{k2}, \ldots, c_{km},]^T \in R^{m\times1}$ is called the center vector of the kth node, σ_k is the bandwidth of the basis function $\phi_k(\bullet)$, and $\|\bullet\|$ denotes the Euclidean distance. For each neuron in the hidden layer, the Euclidean distance between its associated center and the input to the network is computed. The output of the neuron in a hidden layer is a nonlinear function of the distance, and the gaussian function is widely selected as the nonlinear basis function. After computing the output for each neuron, the output of the network is counted as a weighted sum of the hidden layer outputs [52].

In the training procedure, the steepest gradient descent learning process is used to adjust the appropriate settings of the parameters (e.g., weights, centers, and bandwidths), which make the performance of the network mapping optimized. A common optimization criterion is used to minimize the least-mean-square (LMS) between the actual and desired network outputs. LMS error function is,

$$\varphi(r_n) = \frac{1}{2} r_n^2 \tag{10}$$

where $r_n = d(n) - y(n)$ represents the residual error between the desired, $d(n)$, and the actual network outputs $y(n)$. n indicates the index of the series.

The cost function can be defined as an ensemble average errors

$$J(\theta) = E[\varphi(r_n)], \tag{11}$$

where θ is one of the parameter sets of the network.

According to the gradient descent method, the gradient of the cost function $J(\theta)$ needs to be computed. The gradient surface can be estimated by taking the gradient of the instantaneous cost surface. That is, the gradient of $J(\theta)$ is approximated

$$\nabla_\theta J(\theta) = \frac{\partial J(\theta)}{\partial \theta} \approx \frac{\partial \varphi(r_n)}{\partial r_n} \frac{\partial r_n}{\partial \theta} \tag{12}$$

where

$$\frac{\partial \varphi(r_n)}{\partial r_n} = r_n \tag{13}$$

and

$$\frac{\partial r_n}{\partial \theta} = \frac{\partial y}{\partial \theta} \tag{14}$$

The update equation for the network parameters is given by

$$\theta(n+1) = \theta(n) - \mu_\theta \frac{\partial}{\partial \theta} J(\theta) \approx \theta(n) + \mu_\theta r_n \frac{\partial y}{\partial \theta} \tag{15}$$

In the design process of RBFNN, A major challenge is to determine the number of the nodes in the hidden layer. There is no a general method. Considering the cycle influence of wind speed data and the universality of RBFNN model, in this paper, the three-layer RBFNN, which is $24 \times 50 \times 1$ network architecture, is applied to create the wind speed forecasting model in the two regions. The initial weights are randomly selected from [−0.5, 0.5].

3. Experimentation design and results

3.1. Data sets

The Wuwei and Minqin region, which are located in the Gansu Province in China, has abundant wind resources due to its geographical characteristics. The two regions have the potential to be a valuable wind farm site. To investigate actual wind power potential, it is highly desirable to forecast wind speeds in the two regions.

In this paper, the 24 hourly mean data for about one month in Wuwei and Minqin are selected as illustrative examples to evaluate the proposed approach. Fig. 3 shows the hourly wind speed time series in Wuwei and Minqin. In Fig. 3, the above part shows the wind speed data in Wuwei, and the below part shows the wind speed data in Minqin. In the two sites, every site has 744 data, the first 696 values are used as the training sample and the remaining 48 values are used as the testing sample. Table 1 shows the mean, the standard deviation and the maximum and minimum velocities for the two wind speed time series. In this table, it can be observed that the statistical measures of the time series are considerably different among them which are convenient in order to see if the proposed methodology can be applied for different conditions.

Table 1
Statistical measures for the two time series.

Sites	Mean (m/s)	Std. dev. (m/s)	Maximum velocity (m/s)	Minimum velocity (m/s)
Wuwei	2.18	1.38	10.40	0.00
Minqin	3.54	2.45	13.00	0.30

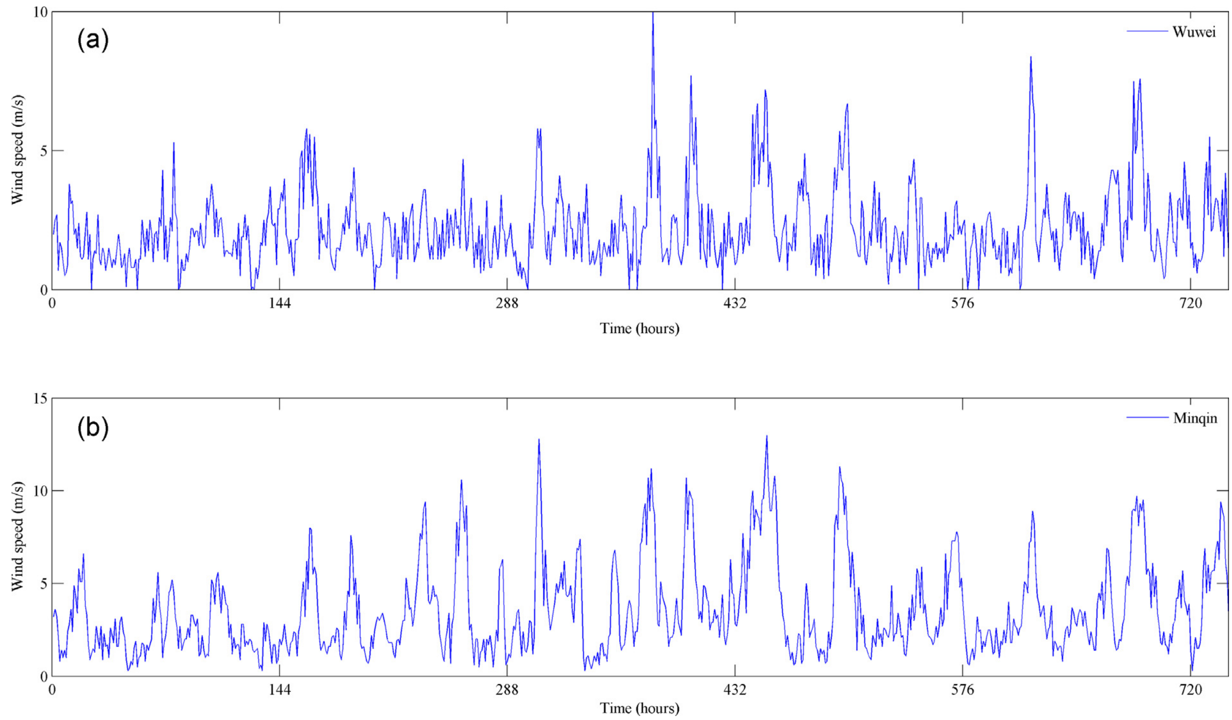

Fig. 3. The mean hourly wind speed data about one month in Wuwei and Minqin.

3.2. Evaluation criteria

In order to determine quantitatively the best model, three forecast error measures are employed for model evaluation and model comparison: the mean absolute error (MAE), root mean-square error (RMSE), and mean absolute percentage error (MAPE). MAE, RMSE, and MAPE are measures of the deviation between actual values and prediction values. The forecasting performance is better when the values of these measures are smaller, and the definitions of these criteria can be found as follows:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|e(t)\right| \tag{16}$$

$$RMSE = \left(\frac{1}{n}\sum_{t=1}^{n}e(t)^2\right)^{1/2} \tag{17}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{e(t)}{y(t)}\right| \tag{18}$$

where $e(t) = y(t) - d(t)$, n is the sample size, $y(t)$ and $d(t)$ are the actual and prediction values at time period t, respectively. Currently, the wind speed forecasting MAPE is ranging from 25% to 40%, which is related to not only the forecasting methods, but also the forecasting horizon and the characteristics of the wind speed in a location. In general, the shorter the forecasting horizon is or the more stable the wind speed variation is, the smaller the forecasting error will be. Otherwise, the forecasting error will be larger [53].

3.3. Wavelet de-noising

Noise information can be obviously seen in Fig. 3. To remove the noise information in the raw data series, the WTT is used for de-noising. In the WTT literature, a lot of wavelet functions are used for de-noising. According to the difference of resolution capability and efficiency, every wavelet function has the corresponding application field. Considering the characteristics of the experimental data, the db3 is applied to remove the noise of the raw experimental data series in this paper. Fig. 4 shows the decomposition process of the raw data series in Wuwei and Minqin. From Fig. 4, it can be seen that the noise information in the two experimental data series is dramatically reduced after the de-noising.

3.4. De-seasonalization process

The seasonal and trend components are the two most commonly encountered phenomena in the wind speed series. From Fig. 4, we can see that the two low frequency signals in Wuwei and Minqin have both seasonal item and trend item. Applying SAM to the low frequency wind speed series in our case study, Fig. 5 shows the de-seasonalization process of the low-frequency component in Wuwei and Minqin. From the Fig. 5, we can see that every low frequency signal is separated to the seasonal component and the trend component, and the left part shows the process in Wuwei and the right part shows the process in Minqin. Through removing the seasonal component from the low frequency signals, the seasonalization is extremely reduced.

3.5. Forecasting results

In this section, the raw wind speed series is predicted. Firstly, the trend component of the raw wind speed series is predicted by RBFNN. Then the forecasting results of the raw wind data series is got by multiplying the seasonal index to the forecasting values of the trend component. The process is described as follows.

3.5.1. Forecasting results of the trend component

RBFNN is used for forecasting the trend component in this section. The RBFNN is constructed by the trend component. Considering the cycle influence of wind speed data, the previous 24 data are fed as input and next data is fed as output, while the multi-step-ahead forecasting is adopted in this paper. In the calculation process, the first 696 data of the trend component are

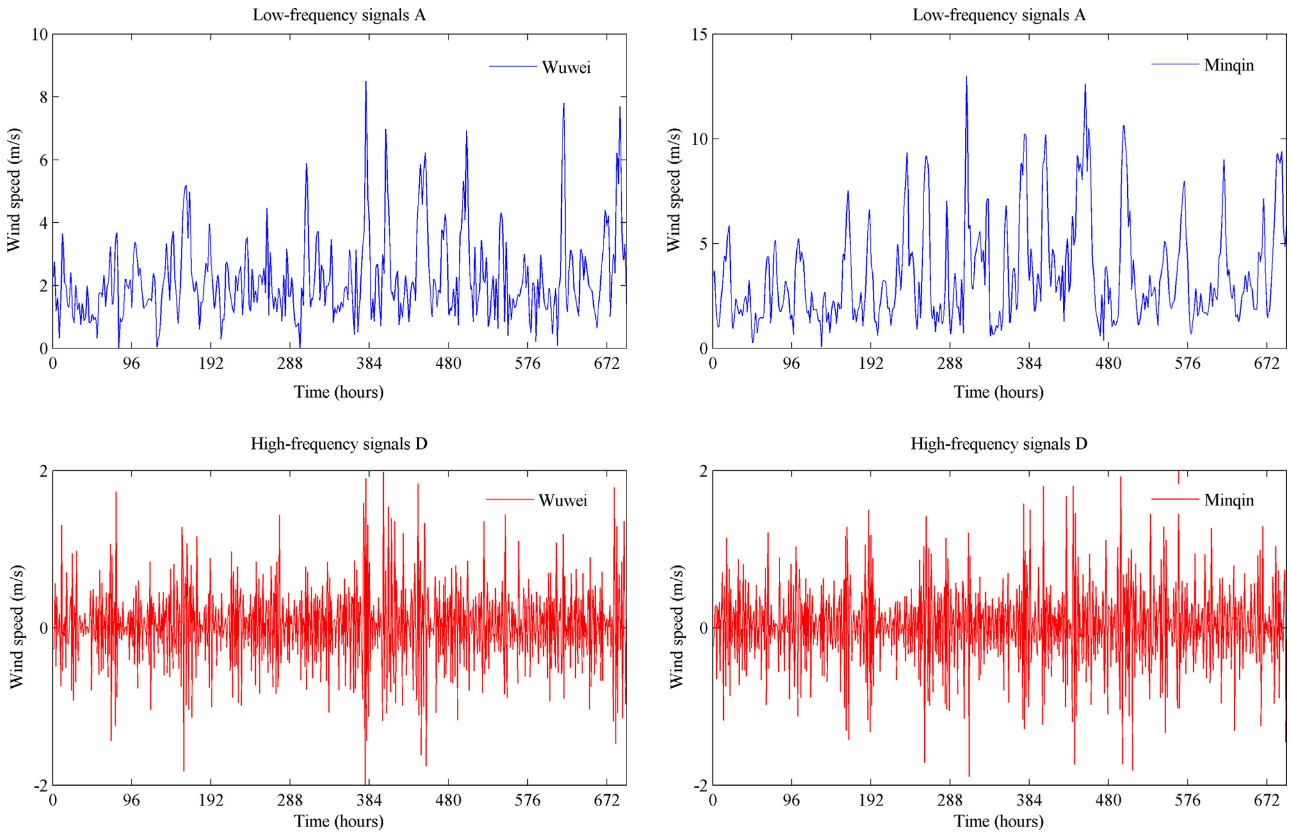

Fig. 4. The WTT decomposition process of the raw data in Wuwei and Minqin.

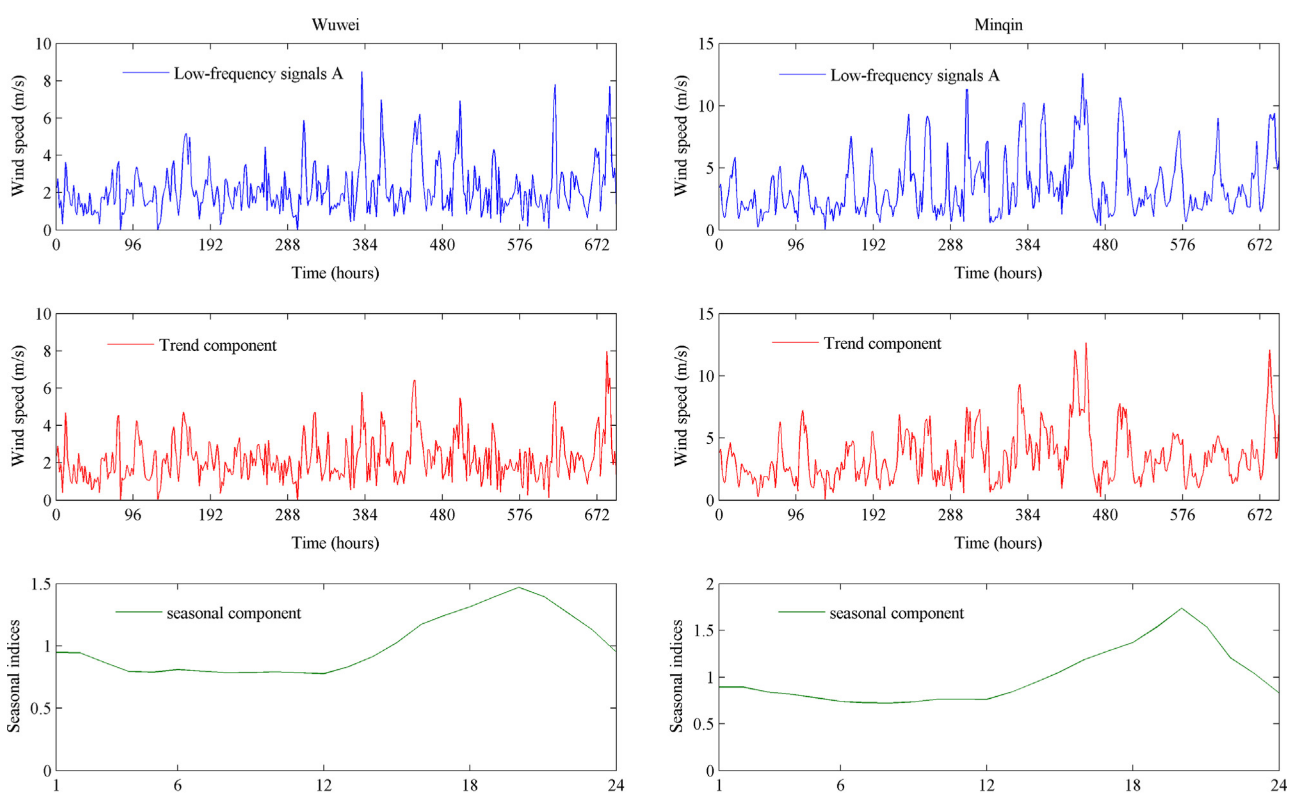

Fig. 5. The de-seasonalization process of the low-frequency component in Wuwei and Minqin.

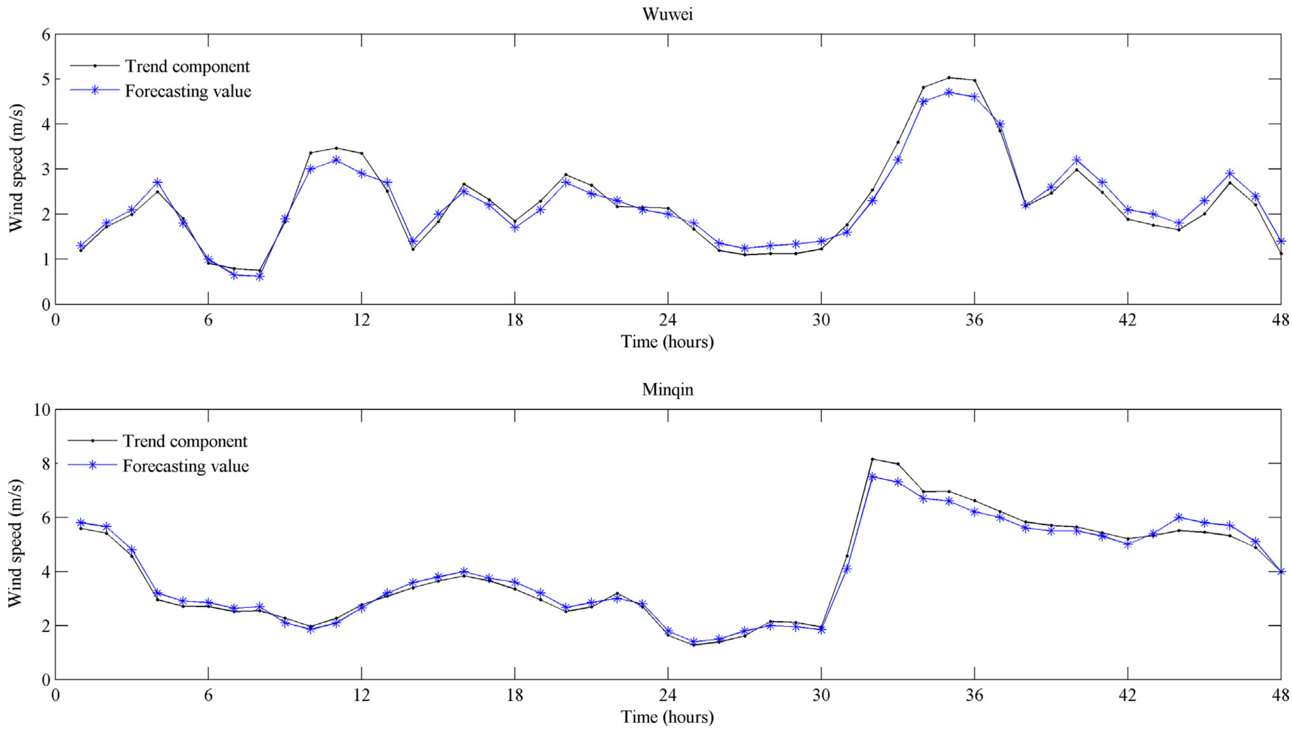

Fig. 6. The forecasting results of the trend component in Wuwei and Minqin.

used as training data, while the remaining 48 data are used to validate the model identified. When RBFNN is used for forecasting, an iterative process is run and the output of the system is fed again as input. So the remaining 48 prediction values are calculated by the iterative process. Fig. 6 shows the forecasting results of the trend component by RBFNN. From Fig. 6, it can be qualitatively observed that the forecasting value of the trend component is close to the real value of the trend component.

3.5.2. *Forecasting results of the raw wind series*

The seasonal index and trend component prediction values are calculated in the previous section. In this section, the final prediction value of the raw wind speed data is calculated by adjusting the seasonal index to the forecasting value of the trend component. Fig. 7 shows the forecasting results of the two raw wind speed series by the proposed approach. In order to validate the prediction capacity of the proposed hybrid

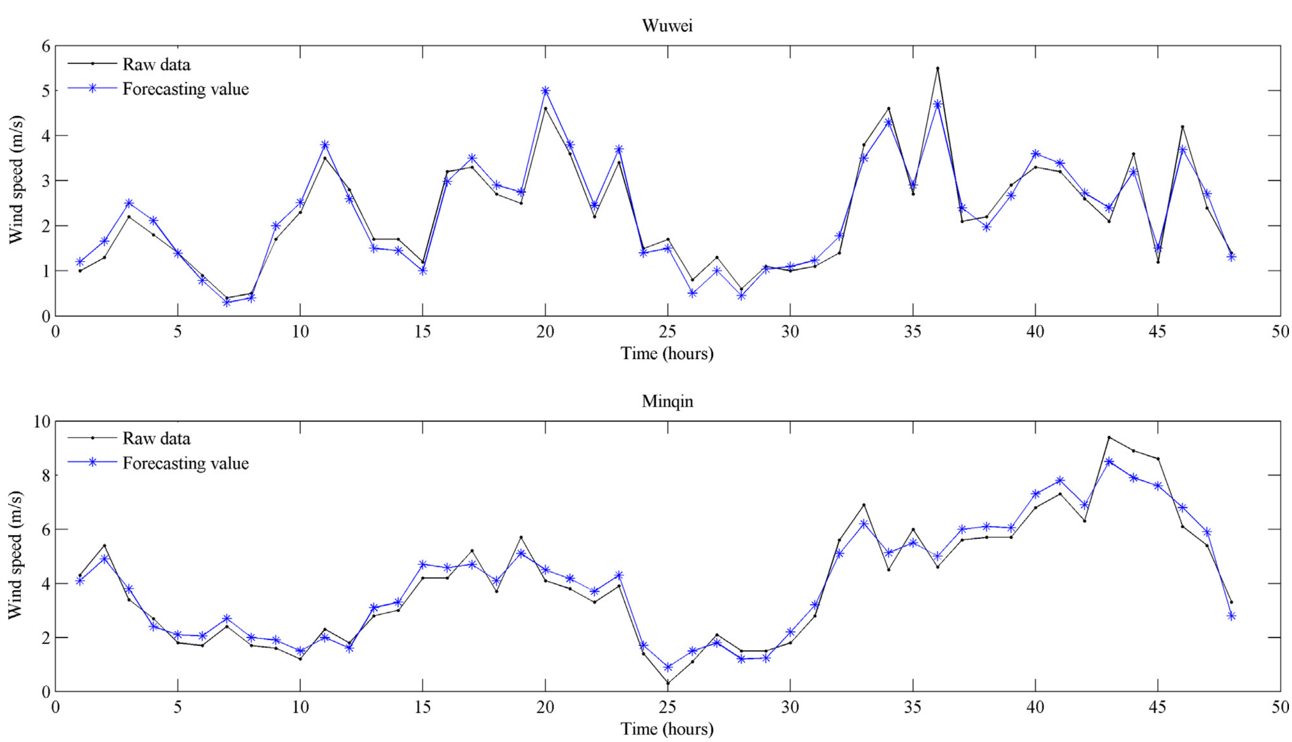

Fig. 7. The forecasting results of the raw wind speed series by the proposed approach.

Table 2
The forecasting results using PM, MLP, RBFNN, SAM-RBFNN, WTT-RBFNN, and WTT–SAM–RBFNN.

Errors	Wuwei						Minqin					
	PM	MLP	RBFNN	SAM–RBFNN	WTT–RBFNN	WTT–SAM–RBFNN	PM	MLP	RBFNN	SAM–RBFNN	WTT–RBFNN	WTT–SAM–RBFNN
MAE	1.21	1.18	1.15	1.06	0.99	0.65	1.42	1.39	1.37	1.16	1.05	0.92
RMSE	1.62	1.55	1.50	1.38	1.25	0.88	1.97	1.93	1.91	1.84	1.78	0.97
MAPE	0.58	0.56	0.53	0.41	0.35	0.23	0.54	0.51	0.48	0.39	0.34	0.28

approach, the model comparisons are given in the next section.

3.6. Model comparisons

The persistence method (PM), also known as a 'Naive Predictor', is generally used as a benchmark for comparing other tools for short-term wind speed forecasting. Wind forecasting methods are usually first tested against the persistence method in order to baseline its performance. In addition, the multilayer perceptron neural network (MLP) is one of the most popular forecasting tools among machine learning methods. To evaluate the performance of the proposed approach, in this paper, the WTT–SAM–RBFNN is compared with PM, MLP, RBFNN, SAM-RBFNN and WTT–RBFNN. The comparison results are shown in Table 2 and it can be clearly seen that the proposed approach consistently has the minimum statistical error of MAE, RMSE and MAPE. The possible reasons are as follows. Firstly, compared with PM, MLP and RBFNN, the WTT–SAM–RBFNN has stronger prediction capacity. It indicates that the WTT–SAM–RBFNN approach can take full advantage of the each individual model that can catch the different information. Secondly, in the model of SAM–RBFNN and WTT–RBFNN, the noise is used in the modeling or seasonal component information is often neglected in these methods, while WTT–SAM–RBFNN considers the actual characteristics of the wind speed data. It is concluded that the proposed approach can improve the forecasting performance and is an effective approach.

4. Conclusions

Taken the highly-noisy and seasonal characteristics of the wind speed series into consideration, a hybrid approach combining WTT, SAM with RBFNN is proposed for wind speed forecasting in this paper. The basic idea of the hybrid model is to remove the noise and consider the seasonal factors in raw data series. As an example, the mean hourly wind speed data about one month in Wuwei and Minqin are applied to validate the forecasting performance of the proposed approach. The results show that the proposed model has smallest error which is better than the others. According to the experiments results, it can be concluded that the proposed hybrid model can improve the forecasting performance in an effective way.

Although we illustrate the effectiveness of using the hybrid model in this paper, there certainly exist limitations. In this study, only the hourly wind speed data are used as illustrative examples to evaluate the performances of the proposed approach. As the hourly data reflect the relatively short-term trend of the wind speed, the predictions generated from the proposed approach provide useful information for the short-term decision-making activities. A possible future research topic is to develop a hybrid model for forecasting with the monthly or even quarterly data. Another direction is to explore the possibility of combining other prediction tools, like support vector regression and multivariate adaptive regression to further improve time-series forecasting.

Acknowledgments

This research was supported by the National Basic Research Program of China '973' Program (Grant No. 2012CB956200), the National Natural Science Foundation of China (Grant No. 71171102), the Opening Fund of Key Laboratory of Land Surface Process and Climate Change in Cold and Arid Regions, CAS, and the Scholarship Award for Excellent Doctoral granted by Lanzhou University.

References

[1] M. Monfared, H. Rastegar, H.M. Kojabadi, A new strategy for wind speed forecasting using artificial intelligent methods, Renewable Energy 34 (2009) 845–848.
[2] J.Y. Zhou, J. Shi, G. Li, Fine tuning support vector machines for short-term wind speed forecasting, Energy Conversion and Management 52 (2011) 1990–1998.
[3] M.S. Miranda, R.W. Dunn, One-hour-ahead wind speed prediction using a bayesian methodology, in: IEEE Power Engineering Society General Meeting, 2006.
[4] X. Wang, G. Sideratos, N. Hatziargyriou, L.H. Tsoukalas, Wind speed forecasting for power system operational planning, in: Proceeding of 8th International Conference on Probabilistic Methods Applied to Power Systems, Iowa State University, Ames, Iowa, September, 2004, pp. 12–16.
[5] L. Ma, S.Y. Luan, C.W. Jiang, H.L. Liu, Y. Zhang, A review on the forecasting of wind speed and generated power, Renewable and Sustainable Energy Reviews 13 (2009) 915–920.
[6] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, E. Feitosa, A review on the young history of the wind power short-term prediction, Renewable and Sustainable Energy Reviews 12 (2008) 1725–1744.
[7] B. Bailey, R. Stewart, Wind forecasting for wind power stations, in: Proceedings of Wind Energy Conversion, Edinburgh, UK, 1987, pp. 265–269.
[8] P. Pinson, G. Kariniotakis, On-line assessment of prediction risk for wind power production forecasts, Wind Energy 7 (2) (2004) 119–132.
[9] R.G. Kavasseri, K. Seetharaman, Day-ahead wind speed forecasting using f-ARIMA models, Renew Energy 34 (5) (2009) 1388–1393.
[10] G.H. Riahy, M. Abedi, Short term wind speed forecasting for wind turbine applications using linear prediction method, Renew Energy 33 (1) (2008) 35–41.
[11] E. Cadenas, W. Rivera, Wind speed forecasting in the south coast of, Oaxaca, Mexico, Renew Energy 32 (12) (2007) 2116–2128.
[12] H. Kantz, D. Holstein, M. Ragwitz, N.K. Vitanov, Markov chain model for turbulent wind speed data, Physica A Statistical Mechanics and its Applications 342 (1–2) (2004) 315–321.
[13] A.D. Sahin, Z. Sen, First-order Markov chain approach to wind speed modelling, Journal of Wind Engineering and Industrial Aerodynamics 89 (3–4) (2001) 263–269.
[14] A. Shamshad, M.A. Bawadi, W.M.A. Wan Hussin, T.A. Majid, S.A.M. Sanusi, First and second order Markov chain models for synthetic generation of wind speed time series, Energy 30 (2005) 693–708.
[15] S.A. Pourmousavi Kani, M.M. Ardehali, Very short-term wind speed prediction: a new artificial neural network–Markov chain model, Energy Conversion and Management 52 (2011) 738–745.
[16] S.A. Pourmousavi Kani, G.H. Riahy, A new ANN-based methodology for very short-term wind speed prediction using Markov chain approach, in: Proceeding of 8th IEEE Annual Electrical Power and Energy Conference, Vancouver, BC, Canada, October 6–7, 2008.
[17] M.A. Mohandes, T.O. Halawani, S. Rehman, A.A. Hussain, Support vector machines for wind speed prediction, Renew Energy 29 (6) (2004) 939–947.
[18] P. Flores, A. Tapia, G. Tapia, Application of a control algorithm for wind speed prediction and active power generation, Renew Energy 30 (4) (2005) 523–536.
[19] T.G. Barbounis, J.B. Theocharis, M.C. Alexiadis, P.S. Dokopoulos, Long-term wind speed and power forecasting using local recurrent neural network models, IEEE Transactions on Energy Conversions 21 (1) (2006) 273–284.
[20] E. Safavieh, A. Jahanbani Ardakani, A. Kashefi Laviani, S.A. Pourmousavi, S.H. Hosseinian, M. Abedi, A new integrated approach for very short-term wind speed prediction using wavelet networks and PSO, in: International Conference on Power Systems, Bangalore, India, December 12–14, 2007.
[21] S.A. Pourmousavi Kani, S.M. Mousavi, A. Kashefi Kaviani, G.H. Riahy, A new integrated approach for very short-term wind speed prediction using linear

regression among ANN and Markov chain, in: International Conference on Power System Analysis, Control and Optimization, Andhra Pradesh, India, 2008.
[22] M. Carolin Mabel, E. Fernandez, Analysis of wind power generation and prediction using ANN: a case study, Renew Energy 33 (5) (2008) 986–992.
[23] S. Salcedo-Sanz, A.M. Perez-Bellido, E.G. Oritz-Garcia, A. Portilla-Figueras, L. Prieto, F. Correoso, Accurate short-term wind speed prediction by exploiting diversity in input data using banks of artificial neural networks, Neurocomputing 72 (4–6) (2009) 1336–1341.
[24] C. Potter, M. Negnevitsky, Very short-term wind forecasting for Tasmanian power generation, IEEE Transactions on Power Systems 21 (2) (2006) 965–972.
[25] T.H.M. El-Fouly, E.F. El-Saadany, M.M.A. Salama, One day ahead prediction of wind speed using annual trends, in: IEEE Power Engineering Society General Meeting, 2006.
[26] N. Qu, H. Mi, B. Wang, Y.L. Ren, Application of GA-RBF networks to the nondestructive determination of active component in pharmaceutical powder by NIR spectroscopy, Journal of the Taiwan Institute of Chemical Engineers 40 (2009) 162–167.
[27] V.M. Rivas, J.J. Merelo, P.A. Castillo, M.G. Arenas, J.G. Castellano, Evolving RBF neural networks for time-series forecasting with EvRBF, Information Sciences 165 (2004) 207–220.
[28] W. Shen, X.P. Guo, C. Wu, D.S. Wu, Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm, Knowledge-Based Systems 24 (2011) 378–385.
[29] C.L. Lin, J.F. Wang, C.Y. Chen, C.W. Chen, C.W. Yen, Improving the generalization performance of RBF neural networks using a linear regression technique, Expert Systems with Applications 36 (2009) 12049–12053.
[30] J. Park, I.W. Sandberg, Universal approximation using radial-basis-function networks, Neural Computation 3 (1991) 246–257.
[31] T. Poggio, F. Girosi, Networks for approximation and learning, Proceedings of the IEEE 78 (1990) 1481–1497.
[32] B. Cannas, A. Fanni, L. See, G. Sias, Data preprocessing for river flow forecasting using neural networks: Wavelet transforms and data partitioning, Physics and Chemistry of the Earth 31 (2006) 1164–1171.
[33] G.P. Zhang, M. Qi, Neural network forecasting for seasonal and trend time series, European Journal of Operational Research 160 (2005) 501–514.
[34] J.C. Goswami, A.K. Chan, Fundamentals of Wavelets: Theory, Algorithms and Applications, Wiley Publishers, 1999, pp. 149–152.
[35] S.T. Li, S.C. Kuo, Knowledge discovery in financial investment for forecasting and trading strategy through wavelet-based SOM networks, Expert Systems with Applications 34 (2008) 935–951.
[36] J.Z. Wang, J.J. Wang, Z.G. Zhang, S.P. Guo, Forecasting stock indices with back propagation neural network, Expert Systems with Applications 38 (2011) 14346–14355.
[37] D.L. Donoho, I.M. Johnstone, Minimax estimation via wavelet shrinkage, Annals of Statistics 26 (3) (1998) 879–921.
[38] S.G. Chang, B. Yu, M. Vetterli, Spatially adaptive wavelet thresholding with context modeling for image denoising, IEEE Transactions on Image Processing 9 (9) (2000) 1522–1531.
[39] J.B. Ramsey, Wavelets in economics and finance: past and future, Studies in Nonlinear Dynamics and Econometrics 6 (3) (2002) 1–27.
[40] T. Li, Q. Li, S. Zhu, M. Ogihara, A survey on wavelet applications in data mining, SIGKDD Explorations 4 (2) (2003) 49–68.
[41] M.S. Hussain, M.B.I. Reaz, F. Mohd-Yasin, M.I. Ibrahimy, Electromyography signal analysis using wavelet transform and higher order statistics to determine muscle contraction, Expert Systems 26 (2009) 35–48.
[42] A. Assume, J. Campbell, F. Murtage, Wavelet-based feature extraction and decomposition strategies for financial forecasting, Journal of Computational Intelligence in Finance 6 (2) (1998) 5–12.
[43] G. Zheng, J.L. Starck, J.G. Campbell, F. Murtagh, The wavelet transform for filtering financial data streams, Journal of Computational Intelligence in Finance 7 (1999) 18–35.
[44] R.M. Rao, A.S. Bopardikar, Wavelet Transforms: Introduction to Theory and Applications, Addison Wesley, Boston, 1998.
[45] R.C. Gonzalez, R.E. Woods, Digital Image Processing, Prentice Hall, Englewood Cliffs, NJ, 2002.
[46] C.J. Lu, Integrating independent component analysis-based denoising scheme with neural network for stock price prediction, Expert Systems with Applications 37 (2010) 7056–7064.
[47] Z.H. Guo, J. Wu, H.Y. Lu, J.Z. Wang, A case study on a hybrid wind speed forecasting method using BP neural network, Knowledge-Based Systems 24 (2011) 1048–1056.
[48] M. Bianchini, P. Frasconi, M. Gori, Learning without local minima in radial basis function networks, IEEE Transactions on Neural Networks 6 (3) (1995) 749–756.
[49] K.P. Sudheer, S.K. Jain, Radial basis function neural network for modeling rating curves, Hydrological Engineering 8 (3) (2004) 161–164.
[50] L. Yu, W. Huang, K.K. Lai, S.Y. Wang, A reliability-based RBF network ensemble model for foreign exchange rates prediction, in: I. King, et al. (Eds.), ICONIP 2006, Part III, Lecture Notes in Computer Science, vol. 4234, 2006, 380–389.
[51] Y.L. Lean, K.K. Lai, S.Y. Wang, Multistage RBF neural network ensemble learning for exchange rates forecasting, Neurocomputing 71 (2008) 3295–3302.
[52] C.C. Lee, Y.C. Chiang, C.Y. Shih, C.L. Tsai, Noisy time series prediction using M-estimator based robust radial basis function neural networks with growing and pruning techniques, Expert Systems with Applications 36 (2009) 4717–4724.
[53] X. Yang, Y. Xiao, S. Chen, Wind speed and generated power forecasting in wind farm in Chinese, Proceedings of the CSEE 25 (2005) 1–5.