DE GRUYTER OPEN

Journal of ELECTRICAL ENGINEERING, VOL 69 (2018), NO2, 148–155

Wind speed prediction with RBF neural network based on PCA and ICA

Yagang Zhang, Chenhong Zhang, Yuan Zhao, Shuang Gao *

Thanks to non-pollution and sustainability of wind energy, it has become the main source of power generation in the new era worldwide. However, the inherent random fluctuation and intermittency of wind power have negative effects on the safe and stable operation of power system and the quality of power. The key solving this problem is to improve the accuracy of wind speed prediction. In the paper, considering the forecasting accuracy is affected by many factors, we propose that, Principal Component Analysis (PCA) is combined with Independent Component Analysis (ICA) to process the sample, which can weaken the mutual interference between the various factors, extract accurately independent component reflected the characteristics of wind farm and achieve the purpose of improving the accuracy of wind speed prediction. At the same time, the adaptive and self-learning ability of neural network is more suitable for wind speed sequence prediction. The prediction results demonstrate that compared with the traditional neural network predicting model (RBF, BP, Elman), this model makes full use of the information provided by varieties of relevant factors, weakens the volatility of wind speed sequence and significantly enhances the short-term wind speed forecasting accuracy. The research work in the paper can help wind farm reasonably arrange the power dispatching plan, reduce the power operation cost and effectively boost the large-scale development and utilization of renewable energy.

K e y w o r d s: renewable energy, wind speed prediction, PCA, ICA, artificial neural network

1 Introduction

Excessive consumption of traditional energy, intensified environmental pollution and frequent extreme weather, such as monster-level hurricanes, the highest temperatures, heat waves *etc*, have urge all countries in the world to provide political impetus to the solution of energy security, ecological environment and abnormal climate through multilateral cooperation. To establish a future supported by clean and sustainable energy, renewable energy, such as solar, wind and geothermal energy, has become the focus of attention of all countries.

Wind energy is the best solution to deal with global climate change and energy crisis because it has the advantage of abundant reserves, pollution and great generation potential. According to data released by the global wind energy council [1], the installed capacity of global wind power in 2016 has increased over 54.6 GW, with a total installed capacity of 486.8 GW, a year-on-year increase of 12.5%. By the end of 2016, China led global wind power market, with the installed capacity of wind power added 23.3GW, with a total installed capacity of 170GW. Each country wind capacity makes great improvement, but the power grid construction is not coordinated with the new energy generation construction, the power grid dispatch capacity is insufficient and the power grid consumption is low, which all make abandoning wind getting worse. Serious wind abandoning phenomenon not only causes the economic loss of wind power enterprises [2], but also greatly weakens the utilization rate of new energy generation, hindering the large-scale development of wind power industry. Reliable wind speed prediction is the key to effectively arrange the power grid dispatch plan and improve the power grid's consumption ability.

The prediction accuracy of wind speed depends on the performance of prediction model, the characteristics of predicted object itself, and data preprocessing techniques. Wind speed prediction is divided into deterministic prediction and uncertainty prediction by the form of prediction results. Deterministic prediction is the single wind speed prediction, uncertainty prediction is estimating the interval of wind speed and fitting the probability distribution of wind speed. At present, most of wind speed prediction is deterministic prediction, including: support vector machine (SVM) method [3, 4], neural network [5-7], time series method [8-10], grey forecast method [11, 12], etc. The inherent random fluctuation of wind speed sequence has great influence on the prediction accuracy, and the larger the fluctuation range is, the worse the prediction precision is. Preprocessing wind speed sequences reduces the complexity of wind speed sequences and the impact of stochastic volatility on prediction accuracy to a certain extent. The popular methods include wavelet decomposition [13-15], principal component analysis [16], independent component analysis [17] and empirical mode decomposition [18].

Wind speed is related to wind direction, temperature, air pressure, humidity and roughness. Because the mea-

* State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing, 102206, China, yagangzhang@ncepu.edu.cn

DOI: 10.2478/jee-2018–0018, Print (till 2015) ISSN 1335-3632, On-line ISSN 1339-309X

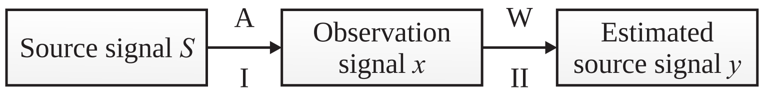

Fig. 1. The process flow chart of ICA

surement of wind speed signal is affected by many factors, we innovatively put forward wind speed predicting model with RBF neural network based on PCA and ICA. At first, the samples are processed by PCA, and we can obtain several comprehensive indexes that affect the wind speed greatly; Secondly, its to make ICA of several comprehensive indexes and wind speed sequence, and get several independent source signals; Then, RBF neural network prediction model is established for each source signal, which will greatly improve the prediction accuracy of a single neural network; Finally, Finally, reconstructing the prediction result of the source signal obtains the final wind speed prediction result. With the conventional prediction methods compared, this method considers the influence of multiple factors on the wind speed sequence, reduces the interaction between various factors and enhance the performance of the RBF neural network effectively.

The structure of this paper is as follows: the second section introduces the theory of PCA and ICA; the third section describes the modeling process of wind speed prediction model with RBF neural network based on PCA and ICA (PCA-ICA-RBF); the fourth section presents the comparison of prediction results and error analysis of each model; the fifth section summarizes the full text and make a prospect.

2 Principal component analysis (PCA) and Independent component analysis (ICA)

PCA is a multivariate statistical analysis technique for data compression and feature extraction that can convert multiple related variables to a handful of uncorrelated principal components. The principal components are usually a linear combination of original variables and can reflect most of original variable information [19]. The data variance reflects the data information. The larger the variance is, the more information is contains. The detailed calculation steps of PCA are as follows [20, 21]:

(i) – Standardize the raw data. Supposing there are m original variables $X_1 X_2 \cdots, X_m$ and n objects. We standardize the original variables to eliminate differences between variables in magnitude and dimension, and get the normalized data matrix.

$$\mathbf{X}_{n\times m} = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \tag{1}$$

(ii) – Establishing the correlation coefficient matrix of standardized data $\mathbf{R}$ and calculating its eigenvalues and eigenvectors

$$\mathbf{R} = (r_{ij})_{m\times m} = \mathbf{X}^{\top}\mathbf{X} \tag{2}$$

We get the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m > 0$ and the corresponding unit eigenvectors

$$\mathbf{u}_1 = \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{m1} \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} u_{12} \\ u_{22} \\ \vdots \\ u_{m2} \end{bmatrix}, \cdots, \mathbf{u}_m = \begin{bmatrix} u_{1m} \\ u_{2m} \\ \vdots \\ u_{mm} \end{bmatrix} \tag{3}$$

(iii) – Ascertain the number of principal components. Selecting the number of principal components depends on the cumulative variance contribution ratio. When the cumulative variance contribution rate of the former p principal components is not less than 95%, it is a good choice to reflect the information of original variable. The variance contribution rate and the cumulative variance contribution rate are respectively

$$\alpha_i = \frac{\lambda_i}{\sum\limits_{j=1}^{m} \lambda_i}, \quad \beta_i = \frac{\sum\limits_{j=1}^{p} \lambda_i}{\sum\limits_{j=1}^{m} \lambda_i} \tag{4}$$

(iv) – The eigenvectors of p principal components are $\mathbf{U}_{m\times p} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p]$ and the matrix of principal components of n samples is

$$\mathbf{Z}_{n\times p} = \mathbf{X}_{n\times m}\mathbf{U}_{m\times p} \tag{5}$$

Independent component analysis (ICA) is a kind of blind source separation (BSS) technique, which actually looks for the directions that can make the data independent in the feature space, and exploring the fluctuation mechanism hidden in the complex phenomena [22, 23]. Supposing $\mathbf{s}(t) = [\mathbf{s}_1(t), \cdots, \mathbf{s}_m(t)]^{\top}$ is dimensional non-Gaussian independent source signal, $\mathbf{x}(t) = [\mathbf{x}_1(t), \cdots, \mathbf{x}_n(t)]^{\top}$ is n dimensional observation signal. For $\mathbf{x}$, the following decomposition can be done,

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{6}$$

Where $\mathbf{A}$ is an $n \times m$ the so called hybrid matrix.

ICA makes statistical independence as the optimization goal. In the absence of both $\mathbf{s}$ and $\mathbf{A}$, we shall find a demixing matrix $\mathbf{W}$, that $\mathbf{x}$ transforms to a new vector $\mathbf{y}(t) = [\mathbf{y}_1(t), \cdots, \mathbf{y}_m(t)]^{\top}$, the component of vector are independent of each other and the optimal approximation of source signals $\mathbf{s}$. That is

$$\mathbf{y} = \mathbf{W}\mathbf{x} \tag{7}$$

The process flow chart of ICA is shown in Fig. 1. The process I – is a mixing process, the process II – is demixing one, $\mathbf{y}$ are the independent components, $\mathbf{W}$ can be obtained and the mixing matrix is approximately $\mathbf{x}\mathbf{W}^{-1}$.

To sum up, in order to find a demixing matrix $\mathbf{W}$, we need to make the following assumptions:

(i) – Components are assumed to be statistically independent;

(ii) – There is only one independent component that does not meet the non-Gaussian features;

(iii) – The unknown mixing matrix is square matrix.

In recent years, the fixed-point algorithm based on negative entropy (Fast ICA) has been widely used, and the algorithm takes the maximum negative entropy as the search direction and gradually extracts independent components. Compared with other independent component analysis algorithms, Fast-ICA has a fast convergence speed, make sure of accuracy and can directly find any non-Gaussian independent components [24]. Negative entropy is defined as follows

$$J(\mathbf{y}) = H(\mathbf{y}_{Gauss}) - H(\mathbf{y}) \tag{8}$$

where, $\mathbf{y}_{\mathrm{Gauss}}$ and $\mathbf{y}$ have the same covariance matrix and both are gaussian vectors, and $H(\mathbf{y})$ is entropy, which is defined as

$$H(\mathbf{y}) = -E\left[\log p_y(\eta)\right] - \int p_y(\eta) \log p_y(\eta) d\eta \tag{9}$$

where, $E\left[\cdot\right]$ is the expectation, $p_y(\eta)$ is the density of a random vector.

Setting $\mathbf{w}$ as the unit vector, which is $||\mathbf{w}|| = 1$, maximizes the non-Gaussian of the corresponding projection $\mathbf{w}^\top\mathbf{x}$. We introduce a nonlinear transformation function G, the equation (8) is approximated as

$$J(\mathbf{y}) \approx \{E\left[G(\mathbf{y})\right] - E\left[G(\mathbf{y}_{Gauss})\right]\}^2 \tag{10}$$

In order to maximize the negative entropy, the following objective function can be obtained

$$\begin{aligned} &\mathbf{w}_i^* = \arg\max\left\{E\left[G(\mathbf{w}_i^\top\mathbf{x})\right] - E\left[G(\mathbf{y}_{Gauss})\right]\right\} \\ &\text{s.t.} \quad \mathbf{R}\left[(\mathbf{w}_i^\top\mathbf{x})(\mathbf{w}_j^\top\mathbf{x})\right]\delta_{ij} \end{aligned} \tag{11}$$

where $\mathbf{R}\left[\cdot\right]$ is the correlation matrix,

$$\delta_{ij} = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases}$$

Usually it is taken

$$G(\mathbf{y}) = -\exp(-\mathbf{y}^2/2) \tag{12}$$

$$g(\mathbf{y}) = G'(\mathbf{y}) = \mathbf{y}\exp(-\mathbf{y}^2/2) \tag{13}$$

In summary, the calculation steps of independent components are summarized as follows [25]:

(i) – Centralize the original data $\mathbf{z}$ so that its mean value is 0;

(ii) – Whitening the centralized data, we can get $\mathbf{x}$. The process of whitening is to find a linear transformation to $\mathbf{x} = \mathbf{V}\mathbf{z}$. The eigenvalues of the covariance matrix $\mathbf{C}_z = E\left[\mathbf{z}\mathbf{z}^\top\right]$ are $d_1, \cdots, d_n$, the corresponding unit eigenvector is $\mathbf{x} = \mathbf{V}\mathbf{z}$. Setting $\mathbf{D} = diag(d_1, \cdots, d_n)$, $\mathbf{E} = (e_1 \cdots, e_n)$, we can get $\mathbf{V} = \mathbf{D}^{-1/2}\mathbf{E}^\top$.

(iii) – Randomly select an initializing vector $\mathbf{w}$ which needs to have the unit norm, that is $\|\mathbf{w}\| = 1$.

(iv) – Increase the number of iterations and update, which is $\mathbf{w} \leftarrow E\left[\mathbf{x}g(\mathbf{w}^\top\mathbf{x})\right] - E\left[g'(\mathbf{w}^\top\mathbf{x})\right]\mathbf{w}$; (v) – Standardize $\mathbf{w}$, namely $\mathbf{w} \leftarrow \mathbf{w}/||\mathbf{w}||$;

(vi) – If $\mathbf{w}$ does not converge, return to the step (4), otherwise it will output $\mathbf{w}_i$ as a column of demixing matrix $\mathbf{W}$. Repeating the above steps one can obtain multiple independent components. It is worth noting that the independent component should be subtracted from the observation signal after extracting an independent component.

3 The modelling process of PCA-ICA-RBF

In order to verify the feasibility of this method, we use real wind speed data from Sotavento wind farm in Galicia, Spain to make modeling experiments. Galicia, located in northwest Spain, faces the Atlantic Ocean, which is a typical Mediterranean climate, hot and dry in summer, and mild and rainy in winter [26]. The wind speed distribution curve in summer and winter is shown in Fig. 2(a)-(b), and it is obvious that the wind speed sequence has high volatility.

Wind speed prediction accuracy is affected by many factors, such as wind direction, air pressure, temperature, humidity, wake and roughness, *etc.* Each variable provides a certain amount of information, but its degree of importance varies. Because the information provided by each variable has a certain overlap in many cases, we use principal component analysis to process each component, and select a few comprehensive variables to explain most of the original data, so as to reduce the prediction complexity.

The influence factors of simulation data including wind direction D, temperature T, air pressure P, specific volume a and specific humidity H, surface roughness R and power E come from Sotavento wind farm in Galicia, Spain. Making use of PCA to analyze the above eight factors, we can get the eigenvalues shown in Fig. 3(a)-(b). The eigenvalues are arranged in descending order, from the fifth principal component, the line segment becomes flat and almost coincides with, which means that eigenvalue change begins to significantly decrease. From the gravel diagram, we can take the first four or the first five principal components, thus combining the cumulative variance contribution ratio to select the main component. As shown in Tab.1, the cumulated variance contribution ratio (CVCR) of the first four principal components in summer and winter were respectively 99.4% and 99.69%, both are larger than 95% and close to 100%. Almost all

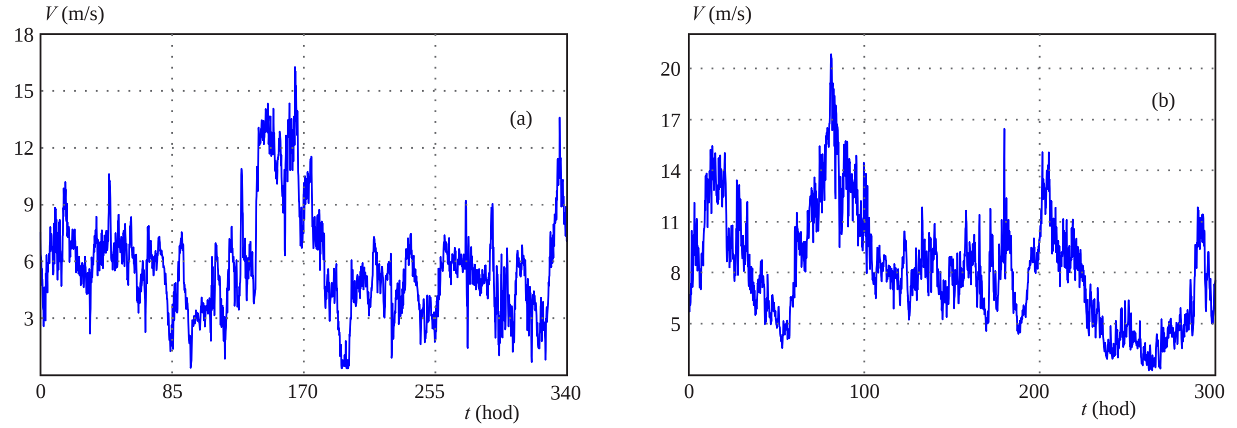

Fig. 2. Wind speed distribution curves in (a)– summer and (b) – winter

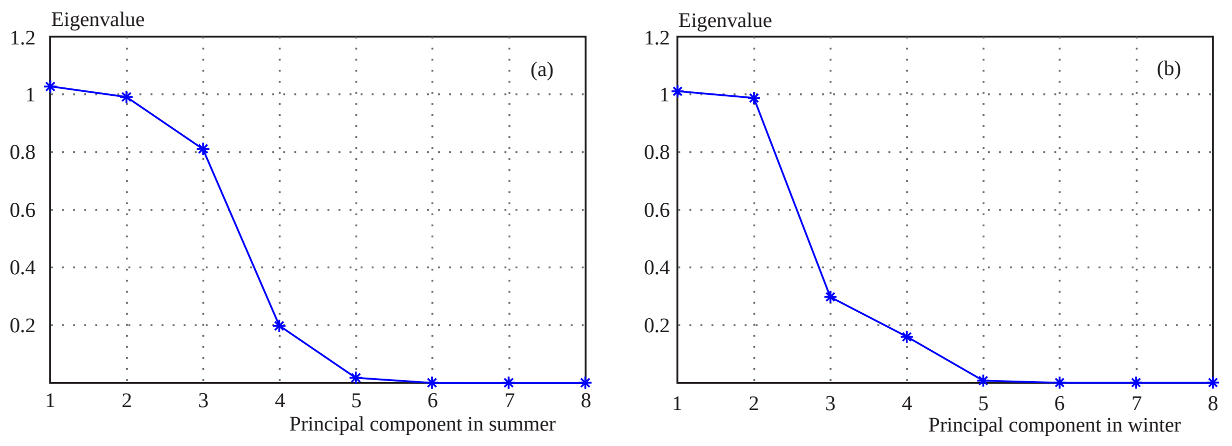

Fig. 3. Eigenvalue of principal components in (a) – summer and (b)– winter

Table 1. The first four eigenvalues, eigenvectors and contribution rates

Temperature	0.0011	0.0001	-0.0018	-0.0003	0.0001	0.0001	-0.0663	0.0058
Air pressure	0.0003	2.7515	-0.001	-0.0010	-2.3741	-5.4046	0.0004	-0.0038
Specific volume	0.0008	0.0001	-0.0007	0.0070	0.0002	0.0002	-0.0067	0.0097
Specific humidity	-0.0125	-0.0013	0.0607	0.0037	-0.0016	0.0017	-0.0422	0.2353
Surface roughness	-0.2456	0.0221	0.6635	0.7048	-0.0378	-0.0087	0.9475	-0.2964
Power	-0.1875	0.0279	0.6789	-0.7081	-0.0199	-0.0007	0.3139	0.9255
Eigenvalue	1.0287	0.9915	0.8113	0.1991	1.0112	0.9876	0.2974	0.1602
Contribution rate	0.3374	0.3252	0.2661	0.0653	0.4104	0.4008	0.1207	0.0650
CVCR	0.3374	0.6626	0.9287	0.994	0.4104	0.8112	0.9319	0.9969

of them explain the total variance, achieve the goal of dimensionality reduction and can well reflect the vast majority of the original variable information.

ICA, as an effective methods looking for the real driving factors behind time series, can transform the observed mixed signals into more structured and more regularized signal to realize the reliable prediction of time series. In this paper, Fast-ICA algorithm is used to estimate five independent components of four principal components and wind speed time series. As shown in Fig.(a)-(b), five independent component components have different wave characteristics in each season. In summer, ICA 1 fluctuates above the x axis with large fluctuations, ICA 2 and ICA 4 fluctuate within the same interval, but volatility characteristics are not the same, ICA3 fluctuation interval is relatively small, while ICA 5 fluctuates below the x axis and the fluctuation interval is large. In winter, the fluctuation interval of each independent component is larger and the change of ICA 1 is smaller.

The principal component analysis can compress multiple influencing factors into several comprehensive components that can describe original data information to reduce dimension and the complexity of data. ICA can decompose non-stationary time series into several inde-

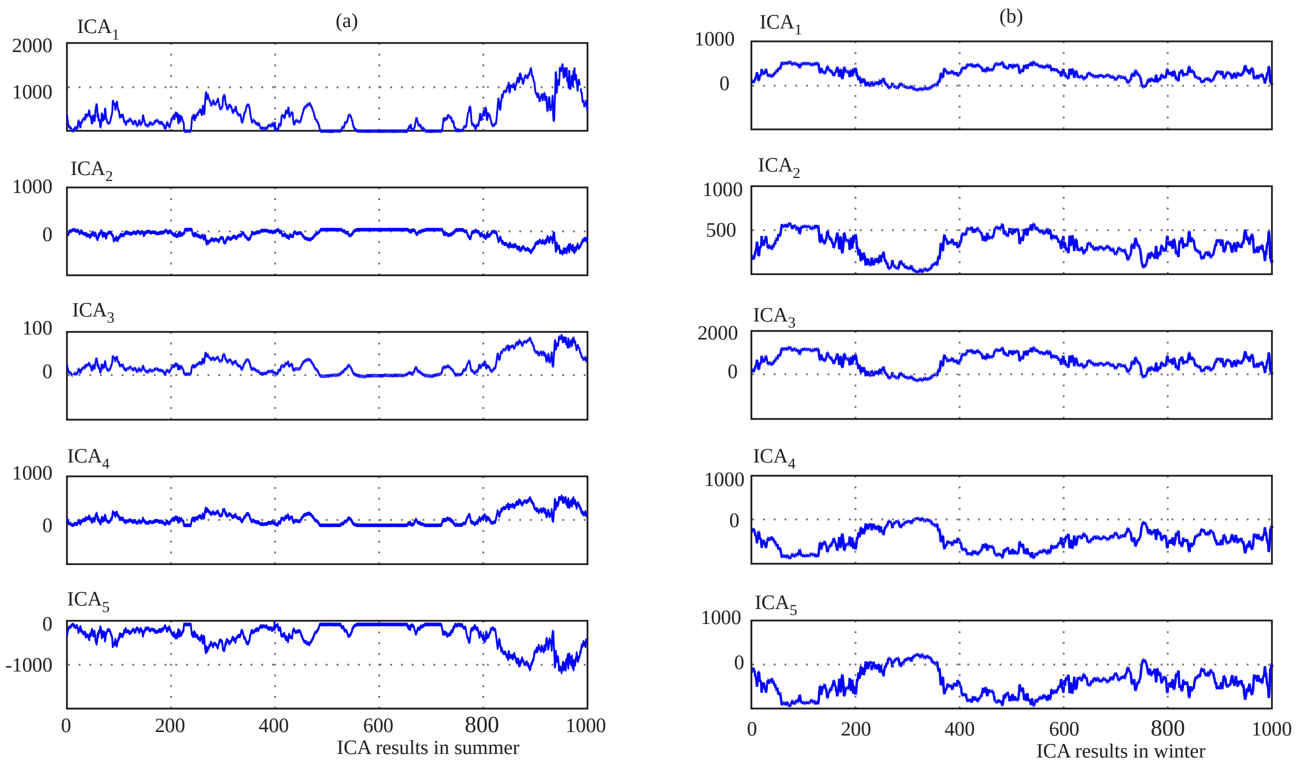

Fig. 4. (a)-(b) Decomposition results of independent components in summer and winter

pendent components with actual meanings, which can help to better grasp the complicated fluctuation mechanism of wind speed series. RBF neural network can realize nonlinear problems prediction well. Integrating the advantage of principal component analysis, independent component analysis and RBF neural network, a PCA-ICA-RBF wind speed prediction model was established. The specific steps are as follows:

Step 1: Principal component analysis of multivariate time series. The PCA model is used to make sine direction and cosine wind direction knowing wind diretion D, temperature T, air pressure P, specific volume a, specific humidity H, surface roughness R and power E, to get several unrelated comprehensive variables describing the overall factors;

Step 2: Independent component analysis of synthetic variables and wind speed time series. The ICA model was used to decompose the data set, and the independent analysis trend obtained by decomposition is used to deeply grasp the fluctuation characteristics of wind speed sequence.

Step 3: Predict each independent component one by one. According to data characteristics of each independent component, we could find the optimal RBF parameters respectively, set up RBF model with best fitting independent components, and predict the independent components.

Step 4: Reconstruct independent component predictions. The predicted values of each independent component are multiplied with the mixed matrix, and the wind speed prediction results are obtained.

Step 5: Error analysis. We make use of mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE) to comprehensively evaluate the prediction results.

$$\begin{aligned} MAE &= \frac{1}{k}\sum_{t=1}^{k}\left|f(t)-\hat{f}(t)\right| \\ RMSE &= \sqrt{\frac{1}{k}\sum_{t=1}^{k}\left(f(t)-\hat{f}(t)\right)^2} \\ MAPE &= \frac{1}{k}\sum_{t=1}^{k}\left|\frac{f(t)-\hat{f}(t)}{f(t)}\right|\times 100 \end{aligned} \tag{14}$$

where, $f(t)$ represents raw data of wind speed, and $\hat{f}t(t)$ represents predicted value of wind speed.

According to the above modeling steps, the modeling flow chart of PCA-ICA-RBF wind speed prediction model is shown in Fig.5.

4 Wind speed prediction results and corresponding error analysis

According to the prediction algorithm in Section 3 and the modeling flow chart, the wind speed prediction model with RBF based on PCA and ICA was established in winter and summer respectively. Making traditional neural network (RBF and BP and ELMAN) as the reference

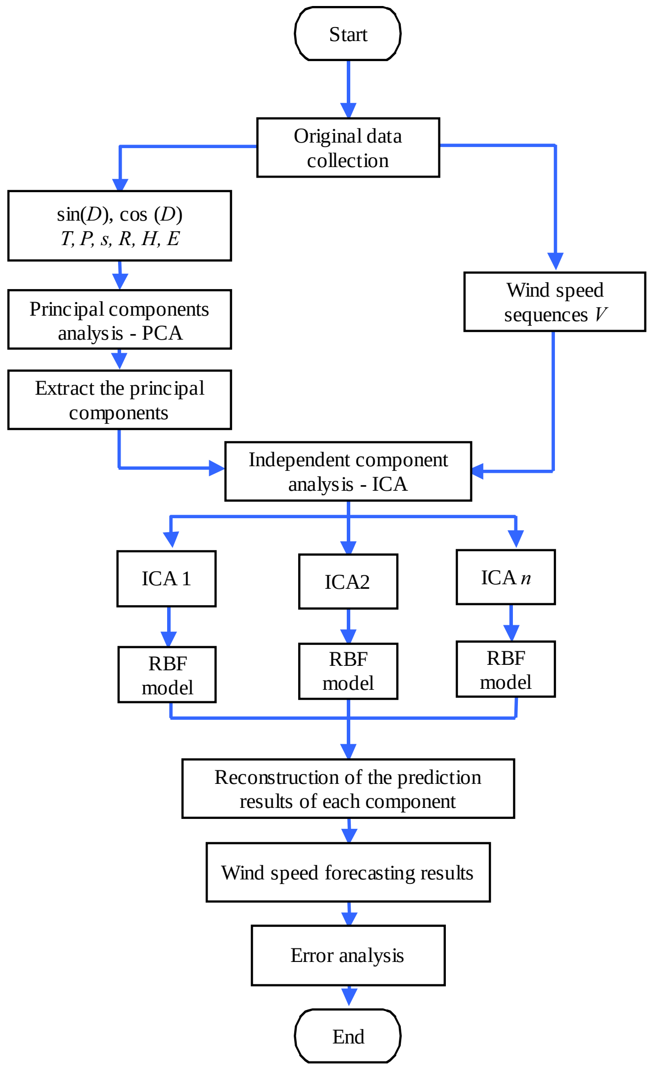

Fig. 5. Flow chart of PCA-ICA-RBF wind speed prediction model

model, we utilize MATLAB to take simulation experiment so as to get wind speed forecasting curve in summer and winter.

Figure 6(a)-(b) describes wind speed prediction curves in summer and winter respectively. The horizontal axis represents the prediction period, and the vertical axis is on behalf of wind speed value. The black curve depicts original wind speed sequence, the blue dotted line depicts wind speed prediction result of Elman model, the green dotted line depicts wind speed prediction result of BP neural network model, the purple dotted line depicts wind speed prediction result of RBF neural network model, the red band point curve depicts wind speed prediction results of PCA-ICA-RBF model. The wind speed prediction results in summer and winter demonstrate that traditional neural network forecasting curve below original wind speed curve can describe the change trend of original wind speed sequence, but the prediction values are greatly deviated from the real wind speed data. Figure (a) is summer wind speed forecasting results for each model. The original wind speed sequence described by the black curve is relatively stable, there are a steep upward trend after the 16th points, three traditional neural networks dont appear good prediction performance, while the red band point curve not only keeps the consistent trend but also closer to the true wind speed curve. Figure (b) describes winter wind speed forecasting results. The black curve describing wind speed sequence has great volatility, which displays that various factors have great influence on wind speed in winter. Meanwhile, there is a sharp increasing tendency and a sharp decline trend after the 9 points, green and blue dotted lines dont reflect the change trend, while purple dotted line describes the change but lag to real data, the red curve is not only in conformity with the trend of original sequence, but also closer to the original wind speed on the numerical data. It fully demonstrates that the PCA-ICA-RBF wind speed prediction model is adaptive to wind speed prediction in different seasons and achieves satisfactory prediction accuracy.

Mean absolute error (MSE), root mean square error (RMSE) and mean absolute percentage error (MAPE) were used to comprehensively evaluate prediction results. Table 2 shows the specific prediction accuracy of the four

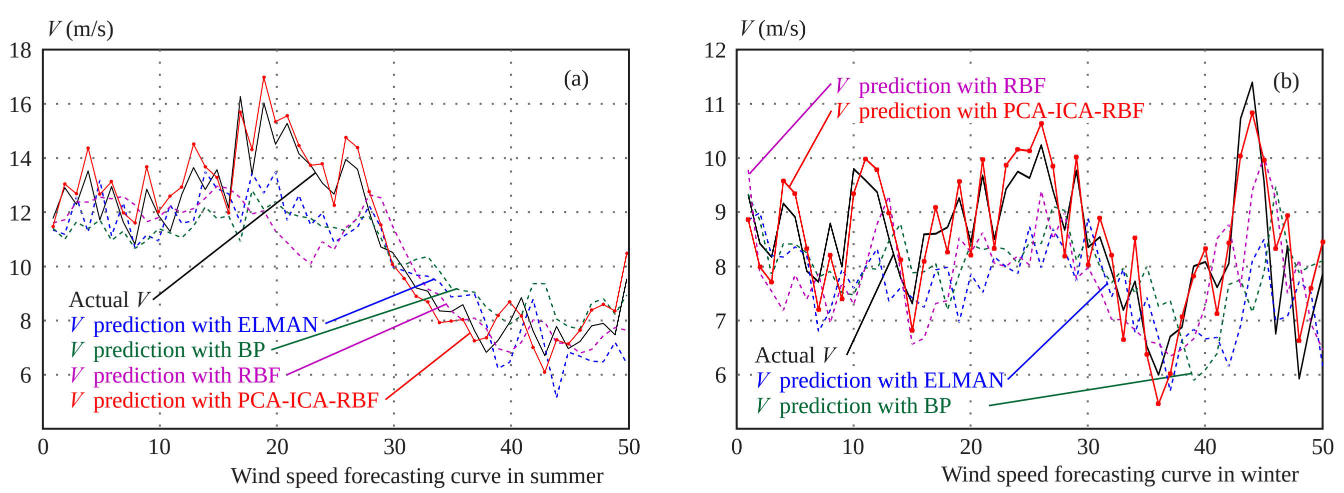

Fig. 6. Wind speed prediction curves of (a) – summer, and (b) – winter season

Table 2. Wind speed prediction model error statistics

Prediction cycle	Prediction model	Error indicator		
		MAE m/s	RMSE m/s	MAPE %
Summer	Elman	1.0945	1.3583	13.1
	BP	1.0177	1.3424	11.94
	RBF	1.0118	1.3289	11.87
	PCA-ICA-RBF	0.4637	0.5099	5.3
Winter	Elman	1.2593	1.6024	14.78
	BP	1.2669	1.6527	13.24
	RBF	1.2348	1.6616	13.11
	PCA-ICA-RBF	0.5651	0.6316	5.61

models. In the winter and summer wind speed prediction results, the differences between MSE and RMSE in three neural networks, Elman, BP and RBF, are very small, and the MAPE error of BP and RBF is about 1.6%, smaller than that of Elman. In summer and winter wind speed prediction results, three error indicators of PCA-ICA-RBF wind speed forecasting model are far less than that of the traditional neural network. Compared with Elman, BP and RBF, MAE is down by at least 54% and RMSE by at least 60%, and MAPE by at least 6.5%. MSE, RMSE and MAPE of the PCA-ICA-RBF wind speed prediction model in summer are respectively 0.4637 m/s, 0.5099 m/s and 5.21%, slightly less than that in winter. Therefore, PCA-ICA-RBF wind speed prediction model had better prediction results for stable wind speed. Through the above analysis, we know that, no matter which season's wind speed sequence is modeled, PCA-ICA-RBF model error indicators are superior to the traditional wind speed prediction model. This again proves the good performance of PCA-ICA-RBF.

5 Conclusions

At present, the worlds new energy technologies have entered highly active period and are promoting the transformation of energy pattern and the market competitiveness of new energy sources at an unprecedented speed. The new energy generation technologies based on wind energy and solar energy have been vigorously supported by governments around the world. The global renewable energy power generation installed capacity is on the rise. It is estimated that by 2040, the proportion of global renewable energy in the overall energy structure will exceed 51 %, this fully shows that the new energy replacing traditional energy has been trending. Wind energy is a clean and stable energy. Wind power generation is universally recognized as a feasible solution to effectively mitigate climate change, improve energy security and promote low-carbon economic growth. However, the stochastic volatility of the wind speed makes the wind power generation changeable and unpredictable, thus affecting the stability of the power system. Reduce adverse influence of wind power generation integrated into the grid is the key to improve wind speed prediction accuracy. Therefore, we use wind farm data in Spain during the winter and summer, take advantage of PCA to extract relevant influencing factors and make use of ICA to deal with influencing factors and wind speed series. Based on the above operation, the hidden information can be mined, and the interference between the time series can be reduced, and the relevant information can be extracted to the maximum extent, the fluctuation of wind speed sequence is eliminated, and the prediction performance of RBF neural network and the wind speed prediction accuracy both are improved. PCA and ICA make conspicuous improvement in RBF neural network prediction, in the next step of research work, we will use other numerical algorithms to carry out simulation experiments and combine the advantages of PCA and ICA to improve wind power prediction accuracy.

Acknowledgements

This research was supported partly by the National Natural Science Foundation of China (51637005), the Fundamental Research Funds for the Central Universities (2017MS166) and the Natural Science Foundation of Hebei Province (G2016502009).

References

[1] "Global wind energy council" (GWEC) [DB/OL].

[2] E. T. Renani, M. F Mohamad Elias and N. A. Rahim, "Using data-driven approach for wind power prediction: A comparative study", *Energy Conversion and Management*, vol. 118, pp. 193-203.June.2016.doi: 10.1016/j.enconman.2016.03.078.

[3] A. Khosravi, R. N. N. Koury, L. Machado and J. J. G. Pabon, "Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system", *Sustainable Energy Technologies and Assessments*, vol. 25, pp. 146-160, February 2018.doi: 10.1016/j.seta.2018.01.001.

[4] S. X. Zhang, Z. P. Zhou, X. M. Chen, Y. Hu and L. D. Yang,"PDHS-SVM: A prediction method for plant DNase I hypersensitive sites based on support vector machine", *Journal of Theoretical Biology*, vol. 426, pp. 126-133, August 2017.doi: 10.1016/j.jtbi.2017.05.030.

[5] F. Bre and J. M. Gimenez, "Prediction of wind pressure coefficients on building surfaces using artificial neural networks", *Energy and Buildings*, vol. 158, pp. 1429-1441, January 2018.doi: 10.1016/j.enbuild.2017.11.045.

[6] J. P. Jeon, C. Kim, B. D. Oh and S. J. Kim, "Prediction of persistent hemodynamic depression after carotid angioplasty and stenting artificial neural network model", *Clinical Neurology and Neurosurgery*, vol. 164, pp. 127-131, December 2017.doi: 10.1016/j.clineuro.2017.12.005.

[7] P. Ramasamy, S. S. Chandel and A. K. Yadav, "Wind speed prediction the mountainous region of India using an artificial neural network model", *Applied Energy*, vol. 80, pp. 338-347, August 2015.doi: 10.1016/j.renene.2015.02.034.

[8] V. Prema and K. Uma Rao, "Development of statistical time series models for solar power prediction", *Renewable Energy*, vol. 83, pp. 100-109, November 2015.doi: 10.1016/j.renene.2015.03.038.

[9] Y. N. Zhao, L. Ye, Z. Li, X. R. Song and Y. S. Lang, "A novel bidirectional mechanism based on time series model for wind power forecasting", *Applied Energy. vol. 177*, pp. 793-803, Mar.2016.doi: 10.1016/j.apenergy.2016.03.096.

[10] Y. G. Zhang, P. H. Wang, P. L. Cheng, and S. Lei, "Wind speed prediction with wavelet time series based on Lorenz disturbance", *Advances Electrical and Computer Engineering*, vol. 17, pp. 107-114, August 2017.doi: 10.1016/j.aece.2017.03.014.

[11] R. Rajesh, "Forecasting supply chain resilience performance using grey prediction", *Electronic Commerce Research and Applications*, vol. 20, pp. 42-58, sep.2016.doi: 10.1016/j.elerap.2016.09.006.

[12] A. Bezuglov and G. Comert, "Short-term freeway traffic parameter prediction: Application of grey system theory models", *Expert Systems with Application*, vol. 62, pp. 284-292, November 2016.doi: 10.1016/j.eswa.2016.06.032.

[13] A. Jackson and B. Turnbull, "Identification of particle-laden flow features from wavelet decomposition", *Physica D: Nonlinear Phenomena*, vol. 361, pp. 12-27, December 2016.doi: 10.1016/j.physd.2017.09.009.

[14] K. Y. Zhang and R. Gencay and M. E. Yazgan, "Application of wavelet decomposition time series forecasting", *Economics Letters*, vol. 158, pp. 41-46, Sep.2016.doi: 10.1016/j.enconlet.2017.06.010.

[15] S.W. Fei and Y. He "Wind speed prediction using the hybrid model of wavelet decomposition and artificial bee colony algorithm-based relevance vector machine", *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 625-631, December 2015, doi:10.1016/j.ijepes.2015.04.019.

[16] V. Gupta and M. Mittal, "KNN and PCA classifier with Autoregressive modelling during different ECG signal interpretation", *Procedia Computer Science*, vol. 125, pp. 18-24, December 2017.doi: 10.1016/j.procs.2017.12.005.

[17] E. P. Duff, A. J. Trachtenberg, C. E. Mackay, M. A. Howard, F. Wilson, S. M. Smith and M. W. Woolrich, "Task-driven ICA feature generation for accurate and interpretable prediction using fMRI", *NeuroImage*, vol. 60, pp. 189-203, Mar.2016.doi: 10.1016/j.procs.2011.12.053.

[18] J. Naik, P. Satapathy and P. K. Dash, "Short-term wind speed and wind power prediction using hybrid empirical mode decomposition and kernel ridge regression", *Applied Soft Computing*, pp. ASOC-4606, Dec. 2017, doi: 10.1016/j.asoc.2017.12.010.

[19] H. Zhang, A. Palazoglu, X. Y. Zhang, W. D. Zhang, Z. M. Zhao and W. S. S. W. L., "Prediction of surface ozone exceedance days using PCA with a non-parametric control limit", *Chemometrics and Intelligent Laboratory Systems*, vol. 133 pp. 42-48, April 2014.doi: 10.1016/j.chemolab.2014.02.005.

[20] S. J. Dong and A. T. H. Luo, "Bearing degradation process prediction based on the PCA and optimized LS-SVM model", *Measurement*, vol. 46 pp. 3143-3152, November 2013.doi: 10.1016/j.mearsurement.2013.02.005.

[21] A. Datteo, F. Luca and G. Busca, "Statistical pattern recognition approach for long-time monitoring of the G. Meazza stadium by means of AR models and PCA", *Engineering Structures*, vol. 153 pp. 317-333, Dec. 2017, doi: 10.1016/j.engstruct.2017.10.022.

[22] J. C. Pereira, J. C. R. Azevedo, H. G. Knapik and H. D. Burrows, "Unsupervised component analysis: PCA, POA, ICA data exploring-connecting the dots", *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 165 pp. 69-84, August 2016.doi: 10.1016/j.saa.2016.03.48.

[23] Y. G. Zhang, P. H. Wang, T. Ni, P. L. Cheng and Shuang Lei, "Wind power prediction based on LS-SVM model with error correction", *Advances Electrical and Computer Engineering*, vol. 17, pp. 3-8, January 2017.doi: 10.1016/j.aece.2017.01.00.

[24] N. Kwak, C. Kim and H. Kim, "Dimensionality reduction based on ICA for regression problems", *Neurocomputing*, vol. 71 pp. 2596-2603, August 2016.doi: 10.1016/j.neucom.2007.11.036.

[25] R. J. Martis, U. R. Acharya and L.C. Min, "ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform", *Biomedical Signal Processing and Control*, vol. 8 pp. 437-448, Sep.2013.doi: 10.1016/j.bspc.2013.01.005.

[26] Y. G. Zhang, J. Y. Yang, K. C. Wang, Z. P. Wang and Y. D. Wang, "Improved wind prediction based on the Lorenz system", *Renewable Energy*, vol. 81, pp. 219-226, Mar.2015.doi: 10.1016/j.renene.2015.03.039.

Received 26 February 2017

Yagang Zhang received his BS degree in Department of Mathematics from Yunnan University, Kunming, Yunnan Province, China in 2001 and the MS degree in Center for Nonlinear Complex Systems from Yunnan University, Kunming, Yunnan Province, China in 2004. He received the PhD degree from North China Electric Power University, China, in 2011. At present, he is an associate professor in the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Baoding, Hebei Province, China. His current research topics include renewable energy source, relay protection of power system and nonlinear complex system theory.

Chenhong Zhang graduated from Handan College, Handan, Hebei Province, China in 2016 and obtained a bachelors degree in mathematics and applied mathematics. At present, she is a Master at Applied Mathematics in North China Electric Power University. Her research area is new energy forecast.

Yuan Zhao graduated from the Hubei Institute of Technology, Xianning, Hubei Province, China in 2017 and obtained a bachelors degree in applied statistics. At present, she is a Master in applied statistics at North China Electric Power University. Her research area is new energy forecast.

Shuang Gao graduated from Xinyang Normal University in Xinyang, Henan Province, China in 2016 and obtained a bachelors degree in mathematics and applied mathematics. At present, she is a Master in applied statistics at North China Electric Power University. Her research area is new energy forecast.