

Ultra-short Term Wind Speed Prediction under Multi-model Structure and Uncertainty Analysis

Yilin Qiao ^a, Yang Hu ^b, Qiong Yang ^c and Dekun Lyu ^d

School of Control and Computer Engineering, North China Electric Power University, Beijing 102200, China.

^aqiaoyilin@aliyun.com, ^bhooyoung@ncepu.edu.cn, ^cqingchunwuhui166@163.com, ^dldkun1993@163.com

Keywords: wind speed prediction; ultra-short term; iterative multi-step prediction; uncertainty analysis; clustering; artificial neural network.

Abstract. In order to improve the forecasting accuracy and reliability of wind power, the precise prediction of wind speed plays a more and more important role. Thus, the ultra-short-term prediction of wind speed and its uncertainty analysis are studied in the paper. Firstly, based on the daily time series of wind speed and its distribution characteristics, the ADAP (adaptive affinity propagation) method is used to execute the clustering of similar days to extract data samples while the PCA (principal component analysis) algorithm is used for dimensionality reduction. Then, the KELM (kernel extreme learning machine) is used to establish prediction models for different wind speed clusters while the PSO (particle swarm optimization) algorithm is applied to optimize the parameters of KELM. To ensure the generalization ability of prediction model, the k-fold cross validation is brought in to obtain final prediction models with stable performance. Subsequently, the iterative multi-step prediction of wind speed is carried out and its uncertainty is analyzed. The minimum standard error model and the mean-square-root standard error model are achieved to describe the comprehensive uncertainty of the prediction methods. Finally, the operation data of wind turbine is used to evaluate effectiveness of the methods. The results show that the prediction accuracy of wind speed is improved by using the combination model with a certain generalization and stability. In addition, the uncertainty models of prediction error are verified to be effective.

1. Introduction

As a clean and renewable energy, wind power generation has been increasingly paid attention among countries all over the world [1]. Wind speed with strong randomness and volatility brings great challenges to large-scale wind power integration. Accurate wind speed prediction is key to lighten impact of grid connected wind power on the power grid [2]. At present, effective wind speed prediction and its uncertainty analysis are a significant way to predict wind speed precisely and reliably, which becomes a prerequisite for operation of grid connected wind farms [3].

It is a research focus about ultra-short term wind speed prediction and its uncertainty analysis based on historical data of wind farm. There are varieties of ultra-short term wind speed forecast, such as continuous method, machine learning method and time series method [4,5,6]. When considering hysteresis and correlation of wind speed's time sequence, it can be used directly to predict wind speed. Because of less uncertain factors, its prediction accuracy is relatively high. But due to unsteady characteristic of wind speed, some problems exist only based on applying time series analysis method, such as low prediction efficiency, limited prediction precision [7]. Some scholars have applied clustering analysis based on similarity theory to divide training samples of prediction model and get similar sample sequence sets of wind speed, which reduces input sample size of prediction model, and improves their similarity [8]. ADAP algorithm, as a kind of unsupervised clustering algorithm, is weak dependence on prior knowledge and outstanding performance in dealing with large sample data sets [9]. Thus, it enables adaptively achieve similar day clustering on the whole year's data of a wind farm. In addition, forecast of wind speed based on time series can be provided with ways of single-step and multi-step, which the latter is of great significance for the prediction of

wind speed. It is direct method and iterative method. Many research results show that iterative one has higher prediction accuracy in multi-step prediction of wind speed [10]. As for a variety of prediction model of wind speed, KELM algorithm with the stronger capability of nonlinear function approximation is selected. Its training efficiency is much higher and its abilities of network generalization and nonlinear fitting are stronger. It is key to select training samples and model parameters [11]. And then prediction uncertainty analysis of wind speed, W. Bo, et.al [12] put forward the two-kind engineering models of comprehensive uncertainty evaluation based on the prediction data in the NWP. They are applied to analyze error distribution between predicted and measured values of wind speed. The confidence interval of prediction error in the NWP and the uncertainty distribution of wind speed forecast are given. The error evaluation model and the uncertainty analysis thought are for references to achieve error distribution of combination model under the different scale-space.

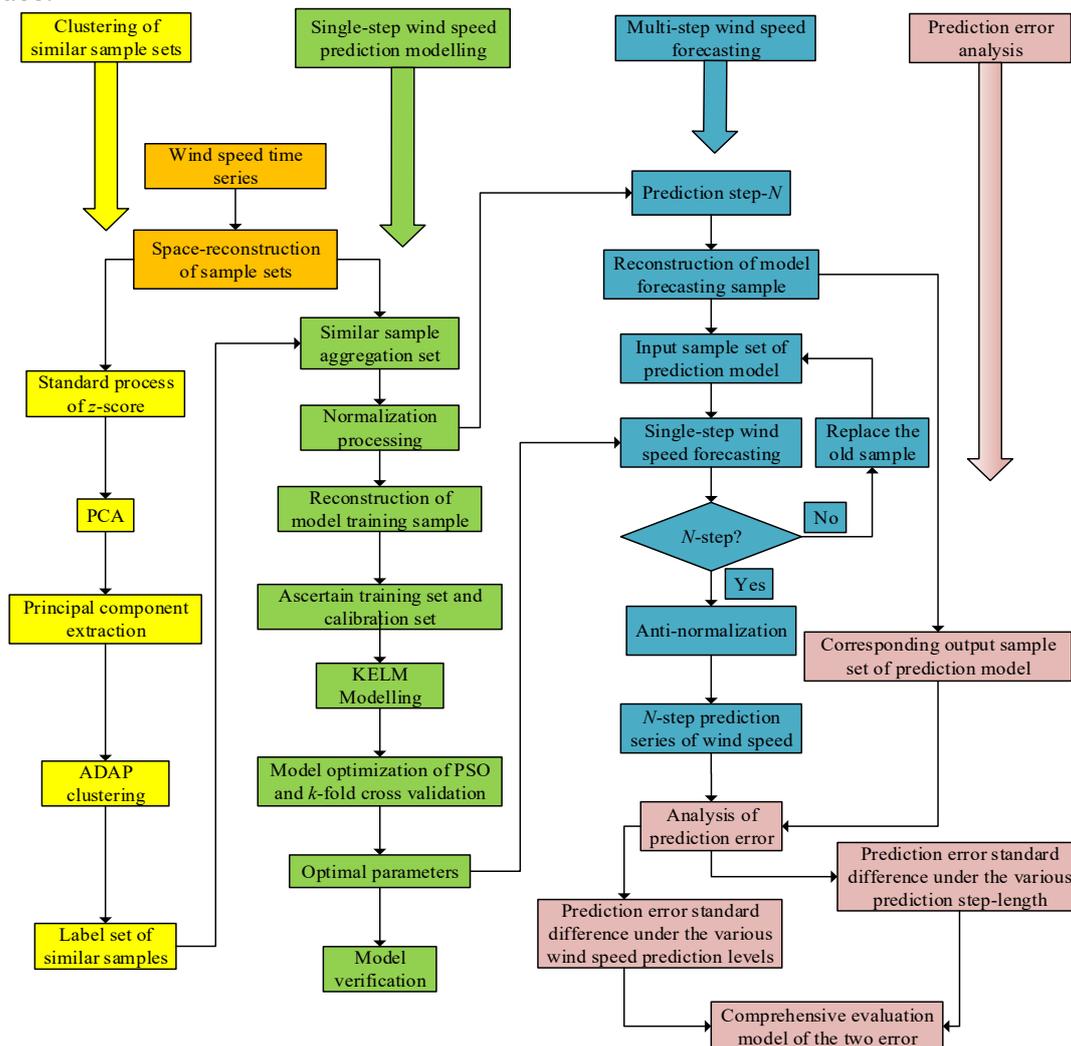


Fig. 1 Flow chart of the whole process

In this paper, PCA is applied to extract main features of original single day wind speed data and to complete dimensionality reduction [13], which provides condensed and reliable wind speed sequences for ADAP algorithm to complement similar day clustering. Then, prediction model of wind speed is built by KELM algorithm. Its parameters are optimized through combination of PSO algorithm [14] and k-fold cross validation method. Finally, iterative method is applied to achieve multi-step wind speed prediction. Prediction error under different step-N or various wind speed levels is analyzed to calculate mean and standard deviation of prediction error sequence. The comprehensive uncertainty evaluation model is constructed to quantitatively describe them. The flow chart of the whole process is shown in Fig.1.

2. Similar data-samples Extraction

2.1 PCA-ADAP Clustering Method.

ADAP is an improved algorithm of Affinity Propagation Clustering(AP). It greatly improves the clustering performance of AP, including adaptive scanning, adaptive damping and adaptive escape. When determining the index of clustering effectiveness evaluation, the Silhouette index, which has good valuation ability, is selected as the evaluation index of the clustering results.

Comprehensively considering training and predicting operational efficiency, applicability and generalization of the prediction model, the similar day set is divided by using ADAP clustering algorithm based on measured 10-minute sampling data of fans. Meanwhile, diversity testing mechanism of sample space is applying to ensure the relatively integrity of structure distribution of wind speed sample set used to train prediction model. Then relativity among training samples in the dataset is raised greatly to achieve reduction of sample size. Make the convergence of model training or prediction better, the operational efficiency and prediction accuracy higher. In addition, the sample diversity ensures a certain generalization ability of prediction model.

According to 10 minutes dense sampling data sequence of wind speed, its dimension is high and distribution presents a certain sparsity and correlation. There is some redundancy and interference noise to some extent. Because of those, the direct implementation of ADAP clustering analysis will cause errors and result in the poor clustering performance. Besides, there are also many problems such as high computational complexity, high dimensional matrix storage and long clustering time. Thus, using PCA algorithm extracts the principal component of wind speed sequence each day. Input matrix dimension of ADAP clustering algorithm will be reduced. That will immensely improve its rate of convergence and efficiency of operation.

2.2 Clustering Result Analysis

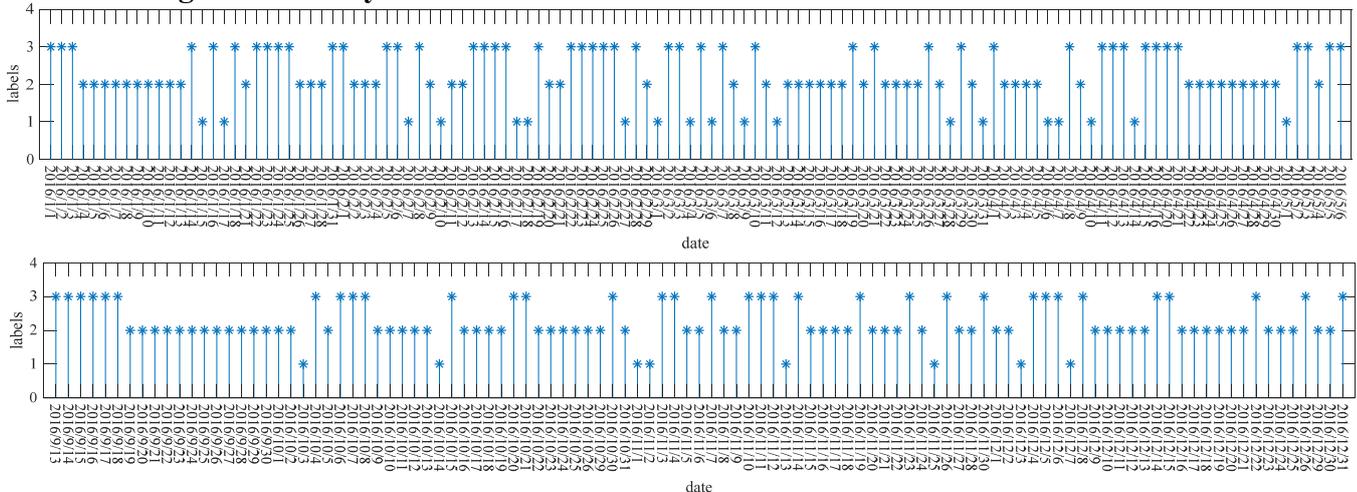


Fig.2 Clustering results of PCA-ADAP

According to the clustering analysis of PCA-ADAP algorithm, wind speed time series is transferred into clustering sample matrix $X_v=[v_{ij}]_{D_n \times 144}$, where D_n is number of effective data days for a year. Before clustering and dimensionality reduction, a data standardized process of z-score is carried out. The clustering result is shown in Fig.2.

3. Iterative multi-step Prediction of Wind Speed

3.1 PSO-KELM Prediction Method.

ELM is a kind of single hidden layer feedforward neural networks with excellent performance. The Kernel function with meeting the condition of Mercer's is applied to replace explicit mapping in the hidden layer of ELM algorithm, which construct KELM algorithm. There exist many kinds of Kernel function, such as Radial Basis Function(RBF), Linear Kernel Function(LKF) and Polynomial Kernel Function(PKF). The output equation of KELM algorithm is as follows:

$$f(x) = \begin{bmatrix} K(x, x_1) \\ \dots \\ K(x, x_p) \end{bmatrix}^T \left(\frac{I}{C} + K(x_i, x_j) \right)^{-1} T \quad (1)$$

where the output weight matrix of it is:

$$\beta = \left(\frac{I}{C} + K(x_i, x_j) \right)^{-1} T \quad (2)$$

Its output value is obtained by only needing to determine the Kernel function form, Kernel parameter(K_para) and penalty coefficient(C).

The prediction model of wind speed, based on KELM algorithm, is trained respectively in the sample dataset. Then wind speed prediction's value is obtained in the various similar day sample sets. Besides, there are some problems about KELM prediction models, such as limitation of generalization ability, lack of prediction precision. By theoretical analysis of ones, its performance of training and prediction is affected by the Kernel function's parameter and penalty coefficient.

A mixed wind speed prediction model of KELM algorithm optimized by PSO algorithm is introduced in this section. The Kernel parameter and penalty coefficient in KELM algorithm are optimized. The basic steps are as follows:

Step 1: determination of model input dimension. The hysteresis and continuity of wind speed variation is considered to determine the reasonable input dimension.

Step 2: initialization of KELM model. According to the distribution characteristics of wind speed sequence and the correlation analysis of sequence values, the appropriate type of kernel function is determined. Then the corresponding boundary conditions of the kernel parameter and penalty coefficient are respectively given.

Step 3: training and optimization of model. Select training sample set. The *k*-fold cross validation is applied to divide sets into subset of *k*-1 and the subset of other one. The former is used to calculate output weight matrix(β) of KELM algorithm, and the latter is used to calculate individual fitness of PSO algorithm, signed as *p_fitness*, whose computational formula is as follows:

$$p_fitness = \frac{1}{k} \sum_{i=1}^k E(t_i, f^{-k(i)}(x_i)) \quad (3)$$

where $f^{-k(i)}(x_i)$ is the prediction model in the *k*-1 sample sets; $E(t_i, f^{-k(i)}(x_i))$ is prediction accuracy in the other *k* sample validation set.

Step 4: model verification. The prediction model after training and optimization is verified in the set of check samples to testify its generalization ability.

Step 5: iterative multi-step prediction. Based on single-step forecasting model of wind speed, its single-step prediction result is used to replace old sample in prediction sample set. That is carried out circularly to achieve *N* steps wind speed prediction.

3.2 Prediction Result and Analysis.

According to PSO-KELM prediction model, it is necessary for wind speed sequence to reconstruct data space. The sequence in a single day is selected with ergodicity to structure wind speed prediction sample pair (X_t, Y_t) , in which $X_t = \{x_{t-m}, x_{t-m+1}, \dots, x_{t-1}\}$, $Y_t = \{y_t, y_{t+1}, \dots, y_{t+n-1}\}$, *m* and *n* are respectively dimension of model input and output. The prediction sequence is shown in Fig.3.

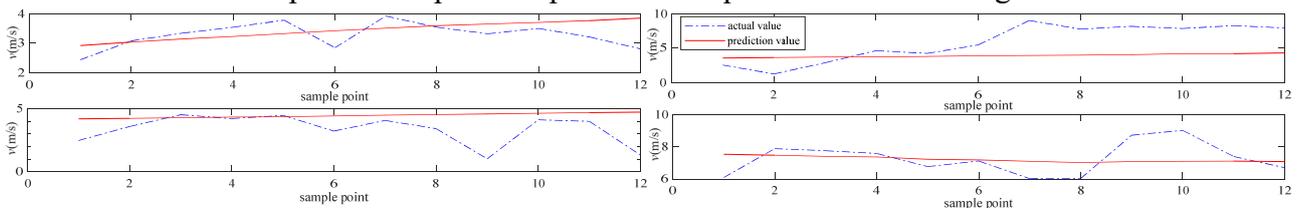


Fig.3 Wind speed prediction series of different time

4. Uncertainty analysis of Predicted Wind Speed

The prediction model of wind speed is applied to achieve wind speed forecasting under the different scale space. First, various prediction step-*N* is selected to complete successively ultra-short

term prediction of wind speed. And error between the prediction value and observation value is also calculated respectively to attain multi-sequence of prediction error(E_T). Second, different forecasting sample sets are constantly choosing to carry out prediction of wind speed under condition of certain fixed prediction step- N_0 . Then based on analysis and integration of prediction result set and corresponding actual wind speed set, wind speed prediction sequence under its different levels, divided by fixed interval- Δv of wind speed, is obtained to calculate corresponding prediction error sequence(E_V). Finally, two-kind engineering models of comprehensive uncertainty evaluation are applied to describe comprehensive forecasting uncertainty of wind speed. Their formulas are shown as follows:

$$\sigma_{\text{mod1}} = \begin{cases} \min(\sigma_T, \sigma_V), v < v_1 \cup v > v_2 \\ \min(\sigma_T, \sigma_V), v_1 \leq v \leq v_2 \end{cases}, \sigma_{\text{mod2}} = \begin{cases} \sqrt{0.5(\sigma_T^2 + \sigma_V^2)}, v < v_1 \cup v > v_2 \\ \sqrt{0.5(\sigma_T^2 + \sigma_V^2)}, v_1 \leq v \leq v_2 \end{cases} \quad (4)$$

where σ_T is the standard deviation of prediction error series under different forecasting step; σ_V is the one of prediction error series under different forecasting levels; v_1 and v_2 are respectively upper and lower demarcation points of standard normal distribution of the error sequences, which is from different scale-space. The result of prediction error in the certain confidence level is shown in Fig.4.

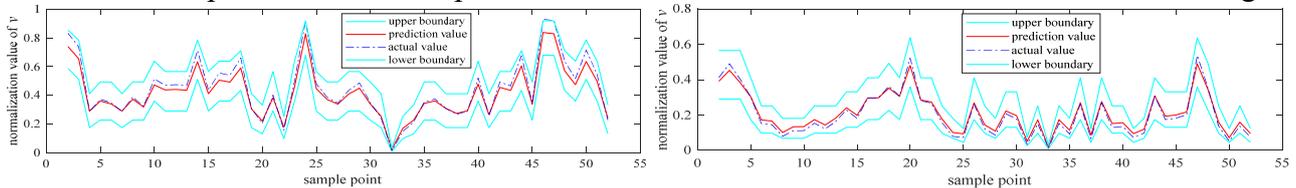


Fig.4 Result of prediction error in the certain confidence level

5. Summary

In this paper, mixed method of PCA-ADAP is used to obtain input sample sequence with high similarity. It can significantly increase validity of input data and improve performance of prediction model. Then training sample set and verification sample set are determined based on the random sampling strategy. Combination algorithm of PSO-KELM is used to construct ultra-short term prediction model. Finally, iterative algorithm of prediction model is formulated, and high-precision multi-step wind speed prediction is realized by circulating one-step prediction model. Through example verification, the following conclusions are drawn.

1) ADAP cluster algorithm realizes similar day set division. It greatly streamlines wind speed prediction modeling training sample capacity, improves similarity between training samples, and provides accurate and reliable data for building KELM prediction models.

2) PSO optimization algorithm is used to optimize key parameters of KELM and determine the optimal prediction model parameters combination, which significantly improves prediction performance of KELM and combines iterative algorithm to achieve accurate multi-step wind speed prediction. In addition, use k-folding cross validation method to enhance prediction model and improve generalization ability and universality of the model.

3) Based on combination prediction model of wind speed, ultra-short prediction of multi-step wind speed is carried out under different step- N and wind speed levels to attain each of forecasting sequences of wind speed, which is used to calculate prediction error series. Then, comprehensive uncertainty evaluation model is applied to integrate the two-kind of prediction error sequences. That achieve a comprehensive quantitative description for prediction error of wind speed.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities (2016MS31, 2018ZD05, 2016MS52), the Project Key Research and Development Plan of Hebei Province(No.18214316D) and the National Natural Science Foundation of China (No. U1766204).

References

- [1]. L. Y. Hua, D. Zhe. "Study on real-time forecasting method of wind speed based on STCP-BP", *AESS*, 37, 4.1009-1016, (2016).
- [2]. Y. X. Ming. "Short term wind speed prediction model based on support vector machine using maximum entropy of chaotic time series", *AESS*, 37, 9. 2173-2179, (2016).
- [3]. J. Wang, F. Zhang. "Hybrid forecasting model-based data mining and genetic algorithm-adaptive PSO: A case study of wind speed time series", *RPG*, 10, 3. 287-298, (2016).
- [4]. C. Y. Zhang, M. Gan. "Predictive deep boltzmann machine for multiperiod wind speed forecasting", *IEEE Transactions on Sustainable Energy*, 6, 4. 1416-1425, (2017).
- [5]. N. C. A. D. Freitas, M. Silva. "Wind speed forecasting: a review", *International Journal of Engineering Science*, 8, 1. (2018).
- [6]. W. Y. Chang. "A literature review of wind forecasting methods", *Journal of Power & Energy Engineering*, 02, 4. 161-168, (2014).
- [7]. H. Q. Zhou, Y. Xue, J. Guo, J. Chen. "Ultra-short-term wind speed forecasting method based on spatial and temporal correlation models", *Journal of Engineering*, (2017).
- [8]. D. Liu, J. Wang, H. Wang. "Short-term wind speed forecasting based on spectral clustering and optimised echo state networks", *Renewable Energy*, 78. 599-608, (2015).
- [9]. J. W. Smith, J. E. Everhart, W. C. Dickson. "Using the ADAP learning algorithm to forecast the onset of diabetes mellitus", *Proc Annu Symp Comput Appl Med Care*, 261–265, (2008).
- [10]. H. Liu, X. Mi. "An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM algorithm", *Renewable Energy*, (2018).
- [11]. Y. Zhang, Y. Xiao, H. Deng. "Modification technology research of short-term wind speed in wind farm based on ELM method", *Meteorological Monthly*, (2016).
- [12]. W. Bo, F. Shuanglei. "Uncertainty evaluation of wind power prediction considering impacts of power curve and predicted wind speed", *Power System Technology*, 38, 2. 463-468, (2014).
- [13]. K. Mahajan, M. R. Vargantwar. "Classification of EEG using PCA, ICA and neural network", *International Journal of Engineering & Advanced Technology iccia*, 1. 80-83, (2011).
- [14]. S. Hua. "Wind speed optimization method of numerical prediction for wind farm Based on Kalman filter method", *Journal of Engineering*, (2017).